UNIVERZITA KARLOVA

Filozofická fakulta

DIPLOMOVÁ PRÁCE

2020 Bc. Klára Divišová

UNIVERZITA KARLOVA

Filozofická fakulta

Ústav anglického jazyka a didaktiky

DIPLOMOVÁ PRÁCE

Bc. Klára Divišová

Víceslovná slovesa v mluvě rodilých a nerodilých mluvčích angličtiny

Multi-word Verbs in Speech of Native and Non-native Speakers of English

Praha 2020

Vedoucí práce: Mgr. Kateřina Vašků, Ph.D.

Poděkování

Ráda bych poděkovala vedoucí práce Mgr. Kateřině Vašků, Ph.D. za trpělivost, vstřícnost a cenné rady, díky kterým bylo možné tuto práci dokončit. Zároveň také děkuji PhDr. Tomáši Gráfovi, Ph.D. za poskytnutí dat k analýze a svým nejbližším za podporu.

Prohlašuji, že jsem diplomovou práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze dne 25. 5. 2020 …..……..…

Klára Divišová

Abstrakt

Diplomová práce se zabývá tématem použití víceslovných sloves (VSS) v promluvách rodilých a nerodilých (českých) mluvčích angličtiny. Konkrétněji je jejím cílem podat jak kvantitativní, tak kvalitativní analýzu tří hlavních VSS kategorií: sloves frázových, předložkových a předložkových frázových. V neposlední řadě práce také shrnuje oblasti a poznatky o VSS v akademických článcích. Motivací k napsání této práce je jedna z oblastí zájmu, tj. vyhýbání se VSS. K analýze byla použita data ze dvou mluvených korpusů: korpusu českých mluvčích LINDSEI_CZ a jeho referenčního korpusu rodilých mluvčích angličtiny, LOCNEC. Analýza ověřuje tři hypotézy: rodilí mluvčí používají více VSS než nerodilí mluvčí, preferovanou kategorií VSS nerodilých mluvčích jsou předložková slovesa a nerodilí mluvčí mají tendenci užívat jistá VSS více (a ve více kontextech) než je obvyklé v kontextu rodilých mluvčích. Výsledky analýzy ukazují, že rodilí mluvčí používají více frázových sloves. Využití předložkových frázových sloves a zejména předložkových sloves je ale u obou skupin spíše srovnatelné. Analýza také ukázala, že nerodilí mluvčí nadužívají (anebo naopak výrazně méně užívají) jistá VSS v porovnání s rodilými mluvčími.

Klíčová slova

víceslovné sloveso, sloveso frázové, sloveso předložkové, sloveso předložkové frázové, frekvence, LINDSEI, LOCNEC

Abstract

The present thesis is concerned with the topic of multi-word verbs (MWV) use in the speech of native and non-native (Czech) speakers of English. More precisely, it aims to give a quantitative as well as qualitative analysis of the use of three main MWV categories: phrasal verbs (PhV), prepositional verbs (PrV) and phrasal-prepositional verbs (PPV). In addition, it summarizes the main research areas in the field of MWV, one of them being the avoidance of MWV by non-native speakers of English, which has been an inspiration for conducting this study. The material comes from two spoken corpora: LINDSEI_CZ corpus of Czech speakers and its referential LOCNEC corpus of English native speakers. The analysis tries to disprove or prove three hypotheses, i.e. non-native speakers' usage of MWV is lower than that of native speakers, prepositional verbs are the favoured MWV by non-native speakers, and non-native speakers overuse certain MWV. The results show that the biggest difference is in the use of PhV as the non-native speakers use significantly fewer PhV than the native speakers; their usage of phrasal-prepositional verbs and especially prepositional verbs is rather comparable to native speakers. Non-native speakers also overuse (and conversely underuse) certain MWV that are far less (or conversely more) frequent in the native corpus.

Key words

multi-word verb, phrasal verb, prepositional verb, phrasal-prepositional verb, frequency, LINDSEI, LOCNEC

Table of Contents

List of Abbreviations and Symbols

* = ungrammatical form

?* = grammatically questionable form

~ = similarity

[VOBJPrt] = structure with a verb followed by an object followed by a particle

[VPrtOBJ] = structure with a verb followed by a particle followed by an object

[V Prt] = structure with a verb followed by a particle

aP = adverbial particle

BNC = British National Corpus

EFL = English as a foreign language

ESL = English as a second language

FC = free combination

ICLE = International Corpus of Learner English

L1 = first language

L2 = second language

LINDSEI = Louvain International Database of Spoken English Interlanguage

LINDSEI_CZ = Czech Louvain International Database of Spoken English Interlanguage

LINDSEI_CZA = Czech Louvain International Database of Spoken English Interlanguage with interviewer's turns only

LINDSEI_CZB = Czech Louvain International Database of Spoken English Interlanguage with students' turns only

LINDSEI_CZAB = Czech Louvain International Database of Spoken English Interlanguage with both student and interviewer's turns

LOCNEC = Louvain Corpus of Native English Conversation

LOCNECA = Louvain Corpus of Native English Conversation with interviewer's turns only

LOCNECB = Louvain Corpus of Native English Conversation with students' turns only

LOCNECAB = Louvain Corpus of Native English Conversation with both student and interviewer's turns

MWV = multi-word verb

NF of Sp = normalized frequency of speakers

NP = noun phrase

Od = direct object

PhV = phrasal verb

PLINSDEI = Polish Louvain International Database of Spoken English Interlanguage

PrV = prepositional verb

PP = prepositional phrase

PPV = phrasal-prepositional verb

[prepO] = prepositional object

S = subject

sb = somebody

Sp = speakers

sth = something

SVA = structure with a subject, a verb and an adverb

V = verb

VSS = víceslovné sloveso

Wh-Q = questions using *wh*-words, i.e. interrogative pronouns

List of Tables and Figures

Figures

Tables

1. INTRODUCTION

Multi-word verbs (MWV) are said to be "a topic of particular importance in English" (Quirk et al., 1985: 1150). Indeed, they are often feared by students of English for their opacity and necessity to be remembered by heart. However, it is a sub-category of verb essential for understanding spoken informal conversation, as well as the works of fiction in which they appear most (Biber et al., 1999: 406). As the academic articles suggest there is not a unified and coherent strategy of presenting them to students and teachers alike. This may be one of the reasons for avoidance of MWV by non-native speakers with various mother tongues which is examined by several authors presented in this thesis, e.g. Dagut & Laufer (1985). Avoidance is also suggested to be caused by unfamiliarity of the structure, indirect interference, arbitrary exposure, and students' lower proficiency in English, e.g. Gilquin, (2015), Yasuda (2010). The studies focused on avoidance as well as learning and teaching has led us to examine the situation in regard to Czech speakers of English.

Thus, this diploma thesis aims to examine the use of MWV, i.e. phrasal, prepositional and phrasal-prepositional verbs, in the speech of Czech speakers of English. The research combines quantitative and qualitative methods. Chapter 2, Theoretical background presents the area of MWV in grammar books as well as in academic articles. The grammar books agree on three main types of MWV and (mostly) on their differentiation: phrasal verbs (i.e. V and adverbial particle combination), prepositional verbs (i.e. V + preposition) and phrasal-prepositional verbs (i.e. V + adverbial particle + preposition) (Biber et al., 1999: 403). Chapter 3, Material and method presents corpora used in the analysis, i.e. LINDSEI and LOCNEC as well as the process of the analysis itself (i.e. tagging the data, their retrieval and categorization) and hypotheses. It is hypothesized that non-native (Czech) speakers of LINDSEI use fewer MWV than the native speakers of English of LOCNEC. Although presupposed to be used less often compared to the native norm, it is expected that the non-native speakers use prepositional verbs the most in their speech (as they are more transparent and thus more accessible than phrasal and also phrasal-prepositional verbs). The non-native speakers are also anticipated to have their favoured MWV that they will overuse as well as use in abnormal contexts, i.e. their "phrasal teddy-bears" (Ellis, 2012: 29; Hasselgård, 2019: 340).

The results of the MWV enquiry are presented in Chapter 4, Analysis. The Analysis is structured into three main areas: analysis of the MWV found in LINDSEI_CZ and LOCNEC respectively and comparison of the results across the corpora. The overall implications as well as comparison of the findings with the theoretical information are presented in Conclusions.

2. THEORETICAL BACKGROUND

The verb category is defined by its grammatical categories (person/number, tense/aspect, mood, voice), formal (non-finite verbs vs finite verbs), syntactic (main verbs vs auxiliary verbs) and semantic features (action vs state). The theoretical background describes a specific verb type, i.e. multi-word verbs (MWV), which are the main concern of this thesis, and their representation in academic articles.

2.1. MULTI-WORD VERBS

The main[1] category of MWV is defined as "verb-particle combinations […] [which] behave as a single lexical unit". (Quirk et al., 1985: 1150); in more detail, it is a "unit which behaves to some extent either lexically or syntactically as a single verb" (ibid.). Thus, although the notion "verb" is usually used for a "morphologically defined word class" (ibid.), here the "verb" is also extended to mean a combination of units that functions as a single verb (ibid., Biber et al., 1999: 403).

When shown on the example of: *We disposed of the problem*, the verb (V) *dispose* may be seen a single unit, "morphologically [as] a verb" (Quirk et al., 1985: 1150), as it can be inflected and conjugated; however, the V + preposition *disposed of* can be also possibly described as a single unit because the combination of *dispose* with *of* part, or e.g. the *by* in *get by* or *in* in *give in* is "morphologically invariable"[2] (ibid.). Thus, because the unit cannot be varied, the whole unit can be classified as MWV. In this invariable unit the first element is always a verb and the follow-up part is called a particle (ibid.), i.e. *dispose* (V) + *of* (particle). The term "particle" is a shielding term for two main classes of words making a verb into MWV; one of them is prepositions, the other is spatial and locative (Biber et al., 1999: 403) adverbials (Quirk et al., 1985: 1151). These two classes of particles, i.e. prepositions and adverbials, are two different, yet coinciding classes, as is shown in Figure 1.

[1] There are also other MWV construction, such as V + V, or V + PP construction (Quirk et al., 1995: 1167). They are discussed briefly in 2.1.3.
[2] In most cases due to semi-opaque exceptions, e.g. *switch/turn on the light*

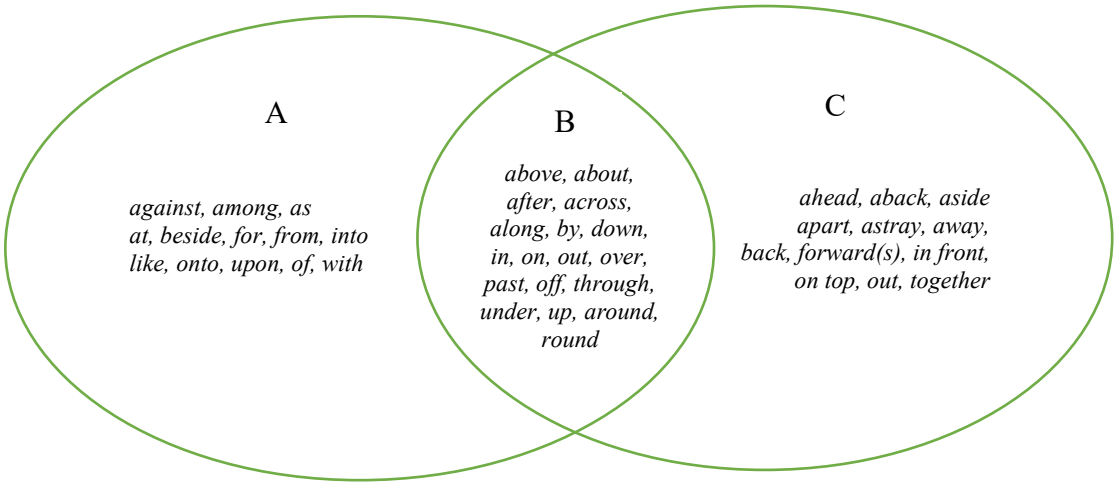

Figure 1 – categories of particles; based on Quirk et al., 1985: 1151

In group A, the particles function only as prepositions; the particles in group C are solely adverbials; however, the particles placed in the group B may function as both prepositions and adverbials, and therefore, they are called "prepositional adverbials" (Quirk et al., 1985: 1151). Alternatively, the complementation of V in MWV is either called an adverbial particle (in the form of either an adverb or adverb/particle, conforming to group C and B) or preposition (group A) (Dušková et al., 2009: 8.51.1-2).

The inflection indicating e.g. tense or person is, as with simple lexical verbs, placed on the verb itself, e.g. *They turned on the light* (ibid.:1153). However, the inflectional suffix is sometimes put onto the particle as a result of "slips of the tongue" (ibid.: 1151): **The editor must do precisely as he see fits* (ibid.) This reanalysis may indicate that MWV is stored in the speaker's head as a unified semantic unit and thus the final element carries the suffix.

Like regular, one-word verbs, MWV may also be categorized on the basis of their semantics. One of such categorizations is to four classes, i.e. activities, accomplishments (expressing continuous tense) (Vendler, 1957: 146), states, and achievements ("lacking continuous tense") (ibid.: 147).

With activities (e.g. *running* or *pushing a cart*) we do not comment on how long the activity takes place and there is no ending point involved in our statement (ibid.: 145), whereas with accomplishment (e.g. *running a mile* or *drawing a circle*) the action has to be completed to have a truth value and to be marked as such (ibid.).

States may last for a longer period of time, e.g. *A loved somebody from* t_1 *to* t_2 (ibid.: 148) whereas achievements happen at a specific time, not for a long period, e.g. *one reaches the hilltop* (ibid.: 146) at one specific moment (ibid.)

Alternative and also broader semantic classification is presented in Biber et al. (1999: 360) where they distinguish seven semantic areas of (both single and multi-word) verbs:

communication, activity, mentality, causativeness, simple occurrence, existence/relationship, and aspect (Biber et al., 1999: 408).

1) Activity verbs

Activity verbs are those verbs that denote some (in most cases) conscious decision to take action, to do something: *move, work* (ibid.: 362), *step up* (ibid.:409), *laugh at* (ibid.: 418). In some cases, the happening is not decided upon, but nevertheless involuntarily happens, mostly when inanimate subjects are involved, e.g. *Compulsory elementary education was working with the vengeance* (ibid.: 362).

2) Communication verbs

Verbs in this category are a special, extended branch of activity verbs that cover the verbs of speaking and writing: *bring up* (ibid.: 409), *respond to* (ibid.: 418).

3) Verbs regarding mentality

Unlike activity and communication verbs, mental verbs are not physical and mostly not based on will; mental verbs denote mental states, such as cognition, e.g. *cope with* (ibid.: 418) *make out* (ibid.: 409) *know,* emotions, e.g. *love, want*; or receptive and perceptive skills: *taste* (ibid.: 362-363), *hear of* (ibid.: 417).

4) Causativeness

Causative and also facilitative verbs bring about a change of condition and thus, a new resulting situation occurs. Prototypical examples are: *call for* (ibid.: 418), *enable, force, help*, or *cause* as in *Still other rules cause the deletion of elements from the structure* (ibid.: 363).

5) Simple occurrence

The main function of these verbs is to "report events" that happen without any personal involvement *run into* (ibid.: 419) or as in *The lights changed* (ibid.: 364); varying based on the interpretation, the lights on the sky has changed, or a person could have pressed a button to change some lights but the action he/she did is the "pressing" not the "changing" itself.

6) Existence/relationship

These verbs reflect (inter)personal relations. Typical existence/relationship verbs are copular verbs, e.g. *All these uses seem natural and serviceable* (ibid.: 364). Other verbs that are assigned to this category are *refer to, rely on*, or *stand for* (ibid.: 419).

7) Aspect

These verbs are connected to an ongoing activity that is happening/was happening and the use of this aspectual verb adds the repetitiveness: *She kept running out of the room* (ibid.: 364), progressiveness: *After another day, he started to recover* (ibid.) or beginning: *begin with* (ibid.: 419).

In conversation, the most frequently used semantic category is the category of activity verbs, followed by mental and existence/relationship verbs (ibid.: 366) which may show the innate ability of humans to share their doings, feelings and thought processes.

MWV can be classified based on several criteria into three major categories: phrasal verbs (PhV), prepositional verbs (PrV) and phrasal prepositional verbs (PPV) (ibid.). Each in turn will be considered here, with further focus being on marginal groups of MWV, i.e. "other MWV constructions" (Biber et al., 1999: 403) and on criteria distinguishing the main groups of MWV.

2.1.1. Phrasal verbs

Phrasal verbs (PhV) are those MWV that have the form of V + adverbial particle (V + aP) (Biber et al., 1999: 403), i.e. particles from group B or C (Quirk et al., 1985: 1151), e.g. *point out, bring up* (Biber et al., 1999: 408). These adverbial particles "all have core spatial or locative meaning" (ibid.: 403). Semantically, V + aP combinations represent, a "single semantic unit" (ibid.); i.e. the meaning of the combination cannot be (mostly) guessed from the two individual parts. This single semantic MWV unit can be recognized by often having a single lexical word synonym, e.g. *put off ~ postpone*, *give in ~ yield* (Dušková et al., 2009: 8.51.1.) or *find out ~ discover* (Biber et al., 1999: 403).

Furthermore, all the V + aP constructions can be put on a scale based on their semantic opacity i.e. idiomatic (opaque) V + aP = PhV (e.g. *give in*) > intensified V+aP (e.g. *drink up*) > non-idiomatic (non-opaque) V + aP (*put back*) (Dušková et al., 2009: 8.51.1). Here, the construction denoting only non-idiomatic V + aP (without any potential figurative meaning) will be disregarded for the core analysis, as those are constructions where the V + aP parts are in free combination (FC) and the meaning is only derivable from the parts.

PhV can be further classified into two distinct subcategories: intransitive phrasal verbs (*come on, shut up*) and transitive phrasal verbs (*bring up, point out*) (ibid.), or in other words, Type I (intransitive) phrasal verbs and Type II (transitive) phrasal verbs (Quirk et al., 1985: 1152-3).

Type I (intransitive) PhV, such as *I hope you'll get by* or *The news made him reel back* (ibid.:1152) do not require object or any other complementation and as such stand on their own, governed by a subject. These PhV can rarely be separated by an additional lexical element: ?* *The news made him reel distractedly back*. FC and PhV can be, in most cases, distinguished by several tests. First, if the elements (V or aP) keep their own distinct meanings, they are not PhV but FC, e.g. *walk past* (ibid.). In this case, one or the other element can also be easily replaced by another one: *walk, swim, jump + over, up, down* (ibid.). Secondly, whereas with

FC it is possible to put *right/straight* between the verb and the particle, e.g. *walk right/straight in* to intensify the action, the use of this modifier is questionable with PhV: **She turned right up at last* (ibid.: 1153). Thirdly, with PhV it is impossible to put the particle at the beginning of a clause with S-V inversion (without inversion when the S is realized by a pronoun): **Up it blew* (ibid.). These tests help to distinguish between PhV and FC, however, they should be taken with precaution, as there are exceptions to the rules presented, e.g. inversion with some PhV may be tolerated (ibid.).

Those PhV that need to take an object complementation are transitive PhV, in other words, Type II PhV, e.g. *He* (S) *can't live down* (PhV) *his past* (Od[3]) (ibid). Some of PhV can behave as both intransitive and transitive, depending on the context. In the case of transitive PhV, it is possible to postpone the particle after the object: *They turned on the light ~ They turned the light on*; if the object is in the form of a personal pronoun it is the only possibility: **They switched on it ~They switched it on* (ibid.: 1154). These two examples of transitive PhV bring about another issue; as was said, PhV are defined as V + aP where neither of the pair can be substituted to still denote the resulting, quite opaque meaning. Here, with *turn on/switch on* it is clear there is an intermediate group to which "some substitutions, but a limited number only can be made into 'semi-idiomatic' example" (ibid.: 1154). Thus, the scale presented with intransitive PhV can be modified to: idiomatic TypeI/II PhV > semi-idiomatic PhV > intensified PhV > non-idiomatic V + aP.

With TypeII PhV, the test of placing the particle before S cannot be used (as transitive verbs overall do not allow for this); however, the *right/straight* intensifiers can be used to distinguish between the PhV and FC. The idiomatic meaning of the V + aP construction also distinguishes the two apart (ibid.). In addition, the passive construction can be also formed with TypeII PhV: *offer was turned down* (PhV) (Dušková et al., 2009: 8.84.13.).

2.1.2. Prepositional verbs

Prepositional verbs (PrV) "[consist] of a lexical verb followed by a preposition with which it is semantically and/or syntactically associated" (ibid.: 1155), e.g. ***Look at** these pictures* or *I **approve of** their action* (ibid.). PrV usually form a clause in two distinct patterns; **1)** NP + verb + preposition + NP, i.e. Type I, without Od; and **2)** NP + verb + NP + preposition + NP, i.e. Type II, with Od (ibid.: 1158, Biber et al., 1999: 413-14).

Type I, i.e. NP + verb + preposition + NP, consists of one NP (after the preposition), which is called in this construction a prepositional object, being the complement of the

[3] S = subject; Od = direct object

preposition: *She* (NP) *looked* (V) *after* (preposition) *her son* (NP) (Quirk et al., 1985: 1156). *Look after* having the near synonym in *to tend* thus should be seen as PrV rather than V + A complementation; this is because although "the verb has a literal use [it still] has a fixed association with the preposition" (ibid.) and for this reason it is classified as PrV and not FC.

Type II PrV form the pattern NP + V + NP + preposition + NP: *But McGaughey bases his predictions on [first-hand experience]* (Biber et al., 1999: 414). The NP after the verb functions as an Od, whereas the one after the preposition is the prepositional object (Quirk et al., 1985: 1158). Type II PrV can be further subclassified into three groups, i.e. Type IIa: *Jenny thanked us for the present*; Type IIb: *Mary took (good) care of the children*; and Type IIc*: I have lost touch with most of the family* (ibid.). However, as Type IIb and Type IIc are governed not only by the V + preposition but also by other element (in this case by more or less fixed NP). Therefore, only the Type IIa will be considered in the analysis as PrV, whereas constructions of Type IIb and IIc will be left to theory as "other constructions".

The Type IIa, V + preposition "form an idiomatic combination" (ibid.: 1159). However, idiomatic here does not mean being a "a single lexical unit" (ibid.) but rather the "verb governing the preposition" (ibid.), in the sense that "the preposition is selected by reason of the verb, rather than by independent semantic choice" (ibid.).

The ability to form passive helps to distinguish between prepositional verb and a verb and preposition in free combination (Dušková et al., 2009: 8.84.12). Whereas the former can be turned into passive, e.g. *this possibility has been reckoned with*, the latter, on the other hand, cannot: **they are often stayed with by their friends* (ibid.). In other cases, the passive may or may not be possible, depending on the abstractness of the verb use, e.g. *The engineers went very carefully into the tunnel > *The tunnel was carefully gone into by the engineers* = concrete use (FC), passive is not possible vs *The engineers went carefully into the problem > The problem was carefully gone into by the engineers* = abstract use (PrV), passive is possible (Quirk et al., 1985: 163).

2.1.3. Phrasal-prepositional verbs

MWV under this heading are those of V + adverbial particle + preposition and are called phrasal-prepositional verbs, i.e. PPV (Biber et al., 1999: 423). They are used mostly in informal contexts (Quirk et al., 1985: 1160). PPV are structured into two patterns; Type I (ibid.): NP + V + aP + preposition + NP (Biber et al., 1999: 423), e.g. *No one has been able to come up with [a product…]* (ibid.) and Type II (Quirk et al., 1985: 1160): NP + V + NP + aP + preposition + NP (Biber et al., 1999: 423), e.g. *I could hand him over to [Sadiq]* (ibid.). Same as PhV, PPV tend to have one-word synonyms: *put up with ~ tolerate* or *look in on ~ visit* (Quirk et al., 1985:

1160). Sentences with PPV can be easily passivized: *such conduct can't be put up with* (Dušková et al., 2009: 8.84.13).

The summarization of the MWV presented above is given in Table 1. It covers only MWV that can be assigned to the three main categories.

			particles		
MWV type	**V**	**Od**	**aP**	**prep**	**[prepO]**
PhV; Type I	Crop	Ø	Up	ø	ø
PhV; Type II	Turn	someone	Down	ø	ø
PrV; Type I	Come	Ø	Ø	across	a problem
PrV; Type II	Take	someone	Ø	for	a fool
PPV; Type I	Come	Ø	Up	with	an answer
PPV; Type II	Put	someone	Up	for	an election

Table 1– summary of the main MWV categories; based on Quirk et al., 1985: 1161

Besides these three major types of MWV, there are other constructions that tend to create somewhat idiomatic units with specific verbs, e.g., verb + PP (*bear in mind*), verb + NP (*take care of*) (Biber et al., 1999: 427/8), i.e. Prv TypeIIb/c; verb + adjective (*lie low*), verb + verb (*get rid of, make do with*), or verb + two prepositions (*compete with N for* N) (Quirk et al., 1985: 1167/8). These 'other' MWV will not be analysed in the Analysis for the reason of space and because they are not exclusively V + particle combinations.

2.1.4. Tests identifying MWV

Tests may be twofold, i.e. semantic (based on the opacity of meaning) and syntactic (movement of the particles, additional, intensifying elements in the phrases, passivization, *wh*-questions, etc.). Some of the tests have been already presented and used, i.e. adverb intensifier insertion, passive formation, prepositional fronting, and particle movement (Biber et al., 1999: 404), e.g. *turn on the light ~ turn the light on* (Quirk et al., 1985: 1153).

Semantic criteria help to distinguish mainly between MWV (especially PhV) and FC; when the meaning can be easily deducible from the construction elements and each of the element keeps their inherent meaning, then we talk about FC; however, when the construction is a single lexical and semantic unit and the meaning is mostly not derivable from the parts, we speak of PhV (Biber et al., 1999: 404). MWV differentiate themselves from FC also by having (in most cases) single-word synonyms (Quirk et al., 1985: 1162).

Syntactic criteria work best for distinguishing PrV and FC (ibid.: 1163), especially for Type I of PrV (as well as PPV). The PrV example *call on* in *He called on a dean* (ibid.: 1163),

meaning ‘to ask some to do something in an official manner’ is clearly idiomatic, we cannot derive the meaning from the parts; however we have to allow for two possible syntactic analysis, either SVA, pointing to FC and SVprepO, pointing to PrV. If the clause can be passivized, i.e. if there is “the possibility of turning the prepositional complement into the subject of the passive sentence” and the preposition is “deferred [in] post verbal position” (ibid.: 1164) we speak of “prepositional passive” (ibid.), e.g. *The dean was called on* vs **Lunch was called after* (ibid.). As with other tests, it is not always true that the prepositional passive indicates only PrV; passive is also possible when the preposition has a locative meaning on its own (ibid.). The ability to turn a sentence into passive cannot be thus taken as a unanimously valid test for PrV; rather it points to the strong connection between the verb and the preposition (ibid.:1165), which is inherently needed for PrV.

The second syntactic test distinguishing between PrV (which is complemented by prepO) and FC (complemented by adverbial PP) are *wh*-questions asking for the prepO (Biber et al., 1999: 405). *Who(m)* and *what* tend to indicate PrV (ibid., Quirk et al., 1985: 1165), e.g. *What are you waiting for* (Biber et al., 1999: 405), whereas when the questions ask about *where*, *when* (ibid.) and *how* or *why* (Quirk et al., 1985: 1165), it indicates FC, e.g. *leave at – When are you leaving?* Lastly, our tendency not to move the preposition away from the V to the beginning of the sentence in relative clauses as well as in *wh*-questions points to PrV, e.g. *She looked after Jim – ?*After whom did she look?* (ibid.: 1166).

Although the tests are helpful, they are not completely reliable with all the cases thus a combination of the tests should be considered in order to create a scale distinguishing PrV from FC. This spectrum is presented in Table 2; it describes whether the passivization is natural or possible, whether *wh*-words *who* and *what* can be used and lastly, if it is impossible to ask about the V + preposition with adverb *wh*-words.

Passivization	Who(m)/what	NO Adv wh-words	**Example**
+	+	+	*The police have asked for details*
+	+	-	*The queen slept in this bed*
-	+	+	*White wine goes with poultry*
-	+	-	*She died of pneumonia*
-	-	+	*His job also comes into the picture*
-	-	-	*She left before noon*

Table 2 – tests distinguishing PrV and FC; based on Quirk et al., 1985: 1166.

Although *come into* is idiomatic in meaning, it cannot be made into passive and we cannot ask about the Od by *who/what wh*-words because the construction is idiomatic only as

a whole (i.e. *come into the picture*), not solely because of the V + preposition. This construction will not be regarded as PrV in the research part due to other elements necessary for creating the idiomatic meaning (ibid.).

PHRASAL VERBS vs PREPOSITIONAL VERBS

It is difficult sometimes to distinguish between PhV Type II, e.g. *He called up the dean* and PrV Type I: *He called on the dean* (Quirk et al, 1985: 1167) when the particle is of group B. There are several tests, both syntactic and phonological that help to decide between the two categories. Similarly, however, they both have one-word synonym, i.e. *called up ~ summon*, and *called on ~ visit* (ibid.).

a) Position of the particle

With PrV it is necessary that the particle stands before the NP and it cannot be separated by such NP from the verb: **They called the dean on*; with PhV the particle can stand before or after the NP without restraint: *They called up the dean ~They called the dean up* (ibid.).

b) Position of a personal pronoun

A special case arises when the NP is realized by a personal pronoun. In this instance, the particle needs to stand before the NP with PrV (*They called on him* vs **They called him on*) and after the NP with PhV (*They called him up* vs **They called up him*) (ibid.).

c) Intensifying adjunct

Without many constrains we insert an intensifier in between the verb and particle with PrV: *They called angrily on the dean*, whereas this is hardly possible with PhV: **They called angrily up the dean* (ibid.).

d) Particle in relative clauses

It is impossible to place the particle of PhV in front of a relative pronoun: **the man up whom they called*, whereas this may be done with a particle in PrV: *the man on whom they called* (ibid.).

e) Particle with *wh*-Q

The same procedure as in point d) is applicable also with *wh*-words; thus, *On which man did they call* (PrV) is possible, but **Up which man did they call* (PhV) is regarded as impossible (ibid.).

f) Stress of particles

Which man did they 'call on vs *Which man did the call 'up*. As marked on the examples, with PrV it is the verb that carries the stress, whereas with PhV it is the particle that is stressed having the nuclear tone when in the final position in a phrase (ibid.).

2.2. MULTI-WORD VERBS RESEARCH

The next section of the theoretical background focuses on the representation of MWV in academic articles. Articles about MWV mostly focus on three main areas[4], i.e. the most frequent MWV, the teaching or learning of MWV, and the tendency to avoid MWV by learners of English.

2.2.1. Most frequent MWV

Several articles focus on the most used MWV, i.e. PhV, namely Trebits (2009), Gardner and Davies (2007) and also Biber et al. (1999) who give a corpus-based analysis of the most frequent MWV.

The study of the most frequent phrasal verbs, or even of the whole multi-word units, has been made easier by the rise of corpus studies (Gardner & Davies, 2007: 340). Gardner & Davies (2007) thus focus their research on the "probabilities of encounter" of phrasal verbs by analysing these constructions in the British National Corpus (BNC) (ibid.). Phrasal verbs are considered an elusive and a "fuzzy grammatical category" (ibid.: 341), which is difficult to describe by linguist, let alone students. The definition of phrasal verbs is thus made as simplistic as possible, stating that phrasal verbs "are all two-part verbs in the BNC consisting in a lexical verb proper followed by an adverbial particle that is either contiguous (adjacent) to that verb or noncontiguous" (ibid.). Both the opaque and transparent verb + particles are considered in the study. The main focus is on 1) the most frequent particles as well as 2) the lexical verbs that function as phrasal verbs and 3) the most frequent phrasal verb combinations (ibid.: 343-344). The most frequent particles are, e.g. *out, up, down, back, off*, etc (ibid.: 346) and they function in majority as particles rather than prepositions (ibid.), still, they all can be categorized into the B group of particles presented in chapter 2.1. This indicates that to rely on "particle identification" only is insufficient and other tests, i.e. "other methods of recognizing PVs" should be also consulted (ibid.: 347).

Most relevantly, the analysis encloses twenty most frequent lexical verbs that combine with particles to form phrasal verbs; these are: *go, come, take, get, set, carry, turn, bring, look, put, pick, make, point, sit, find, give, work, break, hold, move* (ibid.: 349). Not only are these verbs the most frequent verb components in phrasal verbs, they also account for almost 54 % of all the phrasal verbs in the BNC, in other words, "more than half of all the PVs contain a verb from this short list" (ibid.). Further, it is found that these twenty verbs combine with only eight particles, i.e. *out, up, on, back, down, in, over*, and *off*. It is noteworthy that particles seem

[4] Other common areas are e.g. alternative classification of MWV, corpus tagging of MWV, history of MWV

to have preferences for "their" verb, e.g. *on* combines with *go* in 14 743 instances, whereas only once with *point* (ibid.). This highlights the fact that there are not only syntactic but also semantic constrains for the phrasal verb combinations (ibid.: 340), which point to the usefulness of both syntactic and semantic tests introduced earlier in chapter 2.1.4.

While Gardner and Davies (2007), focus on the frequency of phrasal verbs in a large representative corpus, Trebits (2009) studies phrasal verbs in a much smaller corpus of EU English (i.e. in documents of the European Union in English) (Trebits, 2009: 470). Phrasal verbs are described here as "two-word items consisting of a lexical verb and an adverbial particle"[5] (ibid.: 471). The aim of the paper is not only to present the most frequent phrasal verbs in the EU documents (namely booklets, reports, and tests; ibid.: 474) but also to delineate practical implication for teaching phrasal verbs (ibid.: 473). Results show that some of the most frequent verbs in this construction usually take more than one aP to form a phrasal verb; for instance, *take* combines with eight particles, e.g. *take up, take on, take off* (ibid.: 476/7). The most frequent twenty-five phrasal verbs are as follows: *set up*, *set out, base on, carry out, draw up, focus on, lay down, put forward, open up, depend on, make up, report on, find out, call on, move around, take up, follow up, work on, break down, build on, agree on, bring about, go on, point out, speed up* (ibid.: 476). From the semantic point of view, the phrasal verbs tend to have more than one meaning, Trebits (2009) therefore proposes language activities, such as gap-filling or paraphrasing (ibid.: 478) that enable students to practice and use phrasal verbs in different, real-life like contexts using the corpora material (ibid.: 477).

Biber et al. (1999) distinguish between three types of MWV: phrasal verb (PhV), prepositional verbs (PrV) and phrasal-prepositional verbs (PPV). These types have been described in great detail (from the point of view of form and semantics) in chapter 2.1. However, Biber et al. (1999) do not only focus on the description of these MWV, they are also said to give "the best treatment of relative frequency" (Gardner & Davies, 2007: 343) and register use (ibid.). Biber et al. (1999) suggest that PhV mostly appear in fiction and conversation (Biber et al., 1999: 408), for example, *come on* is "the single most common phrasal verb" in conversation (ibid.: 409) and *go on* is the most frequent PhV overall, in all registers (ibid.: 411). Among others that appear in over 20 times per million words in conversation, are e.g. *get out, come over, go off, shut up, sit up* (intransitive verbs denoting activity), *get in, pick up, put on, get back, get off* (transitive verbs denoting activity), *find out, give up* (transitive verbs denoting mental activity), *come off, run out* (occurrence), *turn out* (copular), or *go on* (aspect) (ibid.: 409-410). Other common PhV, i.e. *carry out, take up, take on, set up,* and *point out* occur more

[5] definition is taken from the aforementioned Gardner & Davies (2007) article.

often in written language than in conversation (ibid.: 410). This coincides with findings by Trebits (2009) and Gardner & Davies (2007), as their results prove that these PhV are among the top twenty-five most frequent V + particles in their written corpora (Trebits, 2009: 476, Gardner & Davies: 2007: 358). In addition, verbs that are highly productive on their own, as single lexical verbs, are also those that most readily combine with adverbial particles; those verbs are: *take, get, come, put,* and *go* (Biber et al., 1999: 413). These verbs are also found to commonly form phrasal verbs by Gardner & Davies (2007: 349).

PrV appear most frequently in fiction, followed by conversation (Biber et al., 1999: 415); they are also more frequent in academic prose than PhV, which may indicate that unlike PhV, they are less informal. Overall, they are three to four times more common than PhV (ibid.) *Look at* is the most frequent PrV overall being highly productive in conversation and fiction (ibid.: 416). Other PrV that occur in conversation more than 20 times per million words are, e.g. *look for, go for, go through, play with, do* NP *for, talk to, talk about, speak to, think of, think about, listen to, worry about, know about, look like* and *depend on* (ibid.: 416/17/18). In other registers, e.g. in academic prose: *deal with, be used in, be derived from, refer to, lead to, result in, consist of,* or in fiction: *turn to, stare at* are also very common (ibid.). *In* and *on* are the only productive particles functioning both as prepositions and adverbial particles (ibid.: 422); likewise, only *get* and *go* are the verbs used for forming the most productive PhV as well as PrV. Other than these two, PrV are most readily formed by verbs *look, work, know, hear*, and *use*. However, the V slot in PrV is further taken by various verbs which differ in their semantic sphere; PhV, on the other hand, tend to denote mostly action (ibid.).

PPV are "more common in conversation and fiction, but they are particularly rare in academic prose" (ibid.: 424). Frequent PPV are, e.g. *come out of, get back to* (especially common in the news register), g*o up to, get out of, get on with, get away with, get off at, look forward to, come up with,* and *put up with* (ibid.: 426).

To summarize, there are a few recurrent results and patterns for future analysis. First, there is a clear tendency for some verbs to function as the V in V + particle construction. Secondly, the same can be said about adverbial particles *out, up, down.* Thirdly, as Gardner & Davies (2007) prove and it can be also seen from the other sources, the most frequent phrasal verbs (verb + adverbial particle combination) amount up to 50 % of all the phrasal verbs, indicating that although this field of MWV may be confusing for students, with this information the teachers have the resource to focus on those most probable to appear in a conversation or written texts.

2.2.2. Learning and teaching MWV

For learning and confidently using certain aspect of a language, input and exposure are insufficient for EFL learners as they "are still inaccurate with respect to certain aspects of the L2 grammar" (Nassaji & Tain: 2010: 398). Therefore, it is necessary for student to produce output in order to improve their language capability and skills. Nassaji & Tain (2010) test this hypothesis on the students' usage of PhV; more specifically, they test whether collaborative output raises the chances of learning English phrasal verbs. They focus on whether chosen tasks, i.e. cloze and editing tasks (and also which of the two) are completed more successfully when done as group work and whether students acquire (and retain) more vocabulary knowledge thanks to either of the tasks (ibid.: 401). They use sixteen MWV (meaning "two- or three-word idiomatic expressions, consisting of a verb and a particle or a combination of a particle and a preposition" (ibid.: 402).

The results show that the students are more successful while putting phrasal verbs in the cloze tasks rather than when they are supposed to edit them in the tasks. In addition, both the tasks are completed more correctly, i.e. "more accurate instances of phrasal verbs [were] produced" (ibid.: 406) when the students work together, in pairs. This collaboration is found helpful mainly with the editing tasks. Furthermore, the editing tasks lead to greater acquisition (and retention) of the target MWV when the pre-test (tests done before introducing the target words) and post-tests (after the exposure to the language) are compared (ibid.: 407). Editing task completion also gives rise to more metalanguage between the students, as they are giving each other correcting or confirming feedback. Overall, those working collaboratively succeeded more in the tasks, however the collaboration did not help much with gaining the vocabulary, MWV, knowledge (ibid.: 411).

Unlike Nassaji & Tain (2010), who conduct their research with actual students and are present as observers, Wierszycka (2013) uses the advantage of corpus (namely PLINDSEI, the Polish version of LINDSEI[6]) to study phrasal verbs in the speech of Polish speakers. For her, phrasal verbs (PV) are "a union of a lexical verb and a following particle [that] are to be distinguished from Prepositional Verbs, e.g. *call on* [] and Phrasal-Prepositional Verbs, e.g. *face up to*" (Wierszycka, 2013: 82). Phrasal verbs are distinguished in the article based on syntactic, as well as semantic criteria, ranging from transparent, directional PVs, e.g. *come back*, to semi-transparent, aspectual, e.g. *read through*, to idiomatic, opaque PVs: *come across* (ibid.). Overall, 227 PVs are found in PLINDESI as opposed to 875 in the control corpus of native

[6] Louvain International Database of Spoken English Interlanguage

speakers: LOCNEC[7]; when type/token ration has been considered the result changed to 85 and 274 respectively. This shows that the learners use far fewer PVs than the native speakers.

For the semantic classification of PVs, Wierszycka (2013) first aligns the PVs by particles, as they are more dominant carriers of meaning than the lexical verbs, and secondly, the compositionality of the PVs is taken into account (whether the meaning of the PVs can/cannot be deducible from the parts). Here, again, the categorization should not be seen as finite, but rather as on a pervious scale (ibid.: 88). The results show that the learners "overuse the transparent category (32,1 % vs 24,7 %)" (ibid.: 89). On the other hand, "natives exhibit more types than tokens of PVs" with more idiomatic PVs. In the LOCNEC corpus, when the "idiomaticity grows, so does the use of PVs increase" (ibid.) This has not been observed in the learner corpus, students tended to use more of transparent and mainly semi-transparent PVs (ibid.)[8]. In addition, when the variable "years spent learning English in natural environment and in a classroom" (ibid.: 91) is taken in consideration the expected result (i.e. the more years learners are exposed the more PVs they produce) was not confirmed (ibid.). It is perceived that the quality of exposure rather than the length is the key element for higher use of PVs. Furthermore, both native and non-native speakers agree on shared PVs, which are: *come back, go back, get back, come in, go out, sit down, go on, come on, wake up, find out, make up,* and *show off* (ibid.). In sum, non-native speakers underused PVs in comparison to native speakers; however, at the same time, they use a surplus of semi-transparent PVs.

For Wierszycka's (2013) Polish speakers as well as for Japanese EFL learners in Yasuda (2010) understanding phrasal verbs poses a new challenge, as they do not have this structure in their mother tongues. Also, Japanese students are said to "lack the awareness of the orientational meaning(s) of particles and [fail] to fully understand why one particle is used in a reference to another" (Yasuda, 2010: 251). For this reason, the Japanese students do not perceive the particles as also contributing to the whole construction, although they are the key components on which conceptual and orientational metaphors are projected and their spatial meaning helps with the acquisition of the phrasal verbs (ibid.). The cognitive approach, proposed by Yasuda, view "idiomatic phrases [as] decomposable and analysable, and their meanings are not arbitrary but motivated" (ibid.: 254). For the study, Yasuda (2010) focuses on twenty-one phrasal verbs with particles: *up, down, into, out* and *off,* e.g. *break down, make out, pay off, show up* (ibid.: 257) that are categorized in such orientational metaphors as:

[7] Louvain Corpus of Native English Conversation

[8] The reason for this may be the environment where the data are gathered, i.e. the academic university setting, where the students may feel that using the "colloquial, casual, informal" PVs is inappropriate (Wierszycka, 2013: 89). The same reasoning may be behind the overuse of semi-transparent PVs (ibid.: 90).

CHANGING IS INTO, e.g. *burst into, turn into* or OFF IS STOPING/CANCELLING, e.g. *call off, pay off* (ibid.). The students are divided into two groups, i.e. experimental (i.e. group trained in orientational metaphors) and control group (exposed to phrasal verbs by a traditional method, i.e. translation). Their task is to fill in missing particles (in the phrasal verb-constructions) to thirty clauses. The first fifteen clauses consist of the phrasal verbs the students were exposed to before, whereas the latter part is made out of unfamiliar phrasal verbs (albeit with the chosen particles). As for the first part, both the groups score almost identically well, whereas in the second part: "the experimental group performed significantly better than the control group in the unexposed category of the phrasal verbs" (ibid.: 261).

To conclude, the instruction is not that important for completing the task focused on retrieving the phrasal verbs; however, when new, completely unfamiliar, phrasal verbs come into play, students familiar with the conceptual metaphors tend to produce the target particles more correctly even when they have never encountered the target PhV before (ibid.: 262, 264). The application of orientational metaphors is suggested to be also beneficial for other opaque elements in English, e.g. compound nouns and adjectives (ibid.: 265).

Teaching and learning phrasal verbs by (cognitive) metaphor is also the aim of enquiry for Yang & Hseih (2010). They focus on using cognitive metaphors (working with the notions polysemy and family resemblance in the particles) when teaching phrasal verbs[9] to junior high schoolers in Taiwan (Yang & Hseih, 2010: 2). The authors limited the scope of phrasal verbs to those that consist of particles *up* and *off.* The procedure is essentially identical to Yasuda's (2010): one control group taught by translation and one experimental group taught by metaphors by first explaining that e.g. "the concept finished is often understood as concept *up*" (ibid.: 4) and then by teaching orientational metaphors (ibid.). The testing is done in the same way as in Yasuda's (2010) article, with the only change being the number of task sentences: thirty in Yasuda (2010) vs twelve in Yang & Hseih, 2010: 4). The results show that students in the experimental group perform better overall on the task (ibid.: 5), scoring higher in the unknown phrasal verbs section (71,1 % vs 58,88 %) (ibid.: 6) However, the results disprove the hypothesis stating that conceptual metaphors will help students to retain the phrasal verbs. This is probably due to the fact that the main strategy in Taiwan for learning phrasal verbs is memorization, thus the students perceive the conceptual metaphor technique as an additional burden during their study process which precludes the retention (ibid.: 6). In conclusion, although L1 transfer, traditional way of teaching, and the additional learning burden are the key

[9] Phrasal verbs are defined here as a structure that "combines a verb and invariable particle that function as a single unit both lexically and syntactically" (Liao & Fukuya, 2002: 196) in (Yang & Hseih, 2010: 2).

elements why vocabulary retention is not higher with the experimental group, teaching phrasal verbs by conceptual metaphors may still be beneficial as they help to deal with unknown phrasal verbs, students see the logical, non-arbitrary aspect of the phrasal verbs, and lastly, they allow for phrasal verb categorization (ibid.: 9).

In essence, Yasuda (2010) and Yang & Hseih (2010) build and test their hypotheses on the same grounds. They test whether conceptual, or more precisely orientational metaphors help students of English to learn phrasal verbs, a concept they do not have in their mother tongue. The results show that those introduced to the metaphors can cope better with new, unknown phrasal verbs, especially when the metaphors are introduced at an early stage of learning and are thus not seen as an additional learning burden when compared to the traditional method.

The French are disadvantaged, the same as Japanese and Taiwanese, when it comes to learning phrasal verbs, as they also do not have these constructions in their mother tongue. It is the aim of Gilquin (2015) to study phrasal verbs (both idiomatic and transparent) (ibid.: 4) in the speech and writing of French speakers from the point of view of constructional grammar in terms of frequency and preferences (Gilquin, 2015: 52).

The construction level of analysis focuses on the three layers: 1) PhV in general = superconstruction[10], i.e. frequency of the phrasal verbs overall, 2) structural patterns, focusing on "phrasal verbs that are most distinctively associated with either [VOBJPrt] or [VPrtOBJ][11] (ibid.: 9) and [V Prv] and 3) specific PhV, i.e. most frequent particles and verbs in PhV and the inherent ones to native and non-native speakers. Collostructional analysis was applied to help distinguish between preference for [V Obj Prt] and [V Prt OBJ] as well as between native and non-native use of PhV (ibid.: 9-10).

The analysis shows "higher relative frequency of phrasal verbs", namely twice as many PhV are used in the native data over the non-native, French, data (ibid.: 10). This shows a severe underuse of PhV that is also common with other nationalities and which has led to numerous studies on avoidance (De Cock, 2006; Liao &Fukuya: 2002; Yildiz, 2016, etc.). It is noteworthy that despite Biber et al.'s (1999) findings about PhV being mostly used in conversation, the French speakers use PhV more in writing than in speech (ibid.: 10).

As for the second part, focusing on the three possible structures, the most frequent structure is the intransitive [V Prt] for both groups and in both media corpora (ibid.: 11). [V OBJ Prt] is the least frequently used structure for both the groups; [V Prt OBJ] is used more by the French than by the English natives (ibid.). In speech, the structure [V OBJ Prt] is more

[10] "any combination of a verb and a particle, regardless of the way these two slots are lexically filled" (ibid: 3)
[11] V = verb; OBJ = object; Prt = particle

frequent for both the groups, probably because the speakers rely on shared, live understanding and thus use high number of pronouns in the middle OBJ position (ibid.: 12); on the other hand, when compared from the view point of media, native speakers use more of [V OBJ Prt] in writing, the French rather opt for [V Prt OBJ] with both the media (ibid.). Although the structures [V OBJ Prt] and [V Prt OBJ] are mostly discussed as equivalent, because they are both transitive, Gilquin (2015) shows that when studied in detail, there seem to be a preference of verbs for one or the other constructions (ibid.: 13), e.g. *get back/out/in, put off/in*, and *take in* prefer the [V OBJ Prt] structure (ibid.). Gilquin (2015) finds that this structure is preferred mainly by verbs that are not opaque in meaning and the particles denote spatial aim. Conversely, those favouring [V Prt OBJ] are opaque in meaning and the particles do not retain its spatial or locative meaning. Those PhV are *take on, find out, carry out, give up, work out, point out, take up, bring about,* and *set up* (ibid.: 14)

Lastly, the focus is aimed at the most frequent V and Prt in phrasal verbs. Gilquin (2015) discovers that there is a large discrepancy between the particles studied: some being very low in frequency, e.g. *aboard, under, by, apart*, etc., whereas *out, back, up, on* are in both native and non-native data the most frequent (ibid.: 19). *Back*, consequently, is the most frequent particle for the French speakers (more in speech than in writing), probably because of the mental correspondence with the prefix *re-* on verbs in French (*go back ~ retourner*). The most frequent verbs: *go, come, take, bring, put, carry, turn*, and *give* quite coincide for both the groups (ibid.: 20), whereas *move, end, work* appear in the native group, and *keep, live, build* in the non-native one (ibid.). Thanks to colloconstructional method Gilquin (2015) also presents the most frequent phrasal verbs; for native speakers, the most frequent PhV based on this method of analysis are, e.g. *work out, get on, fit in, start off, come in, come out, get away* and for non-native speakers: *bring together, live together, come back, keep on, sum up, switch on* (ibid.: 22-23). When the data are compared, the non-native speakers rather opt for non-idiomatic PhV (ibid.: 23).

To briefly summarize, the French use significantly less phrasal verbs than native speakers, and when using transitive ones, they opt for the [V Prt OBJ] structure, unlike the native speakers. Although it is found the French use less PhV, they have the same order of preference of PhV with native speakers i.e. [V Prt], [V OBJ Prt] least frequent overall and in writing, and [V Prt OBJ] least frequent in speech. (ibid.: 25). Finally, the French overuse the particle *together* and have a preference for the PhV with particle *back* due to their association with its French counterpart.

Ellis (2012) who focuses on the way a language is learned, supports the idea that second language learning is mainly based on learning "formulaic sequences and their interpretations" (Ellis, 2012: 17). Although he does not specifically mention MWV or phrasal verbs in his work, he presents an interesting concept of phrasal "teddy-bears" which are described as "formulaic sequences […] learnable by dint of being highly frequent and prototypical in their functionality" (ibid.: 29). It is clear that phrasal "teddy-bears" are something the learners overtly rely on. To an extent, this may be also applied to MWV, in a way that students overuse a certain type of PhV or PrV that may not be as frequent in the native mental lexicon, as they are confident with its use and function, considering it also heavily used in native-like setting. In addition, Hasselgård (2019) in her study expands on the notion of these heavily overused sequences, focusing this time on lexical bundles and their under/overuse by non-native speakers. These familiar structures are called "phraseological teddy-bears" (Hasselgård, 2019: 340). Unlike Ellis (2012), who connects the phenomenon mainly to high overuse, Hasselgård (2019) sees her phraseological teddy-bears as a "multi-word unit that learners use more frequently and in more contexts than native speakers do" (ibid.). Therefore, not only are these multi-word units overused in learners' production, they are also overgeneralized and thus used in context where native speakers would opt for different structure (ibid.). In her analysis she does not rely on only raw frequencies but also on dispersion of lexical bundles in the (two) corpora used, meaning that the focus is not only on the quantity of times a certain bundle occurs but also on the distribution in the corpora itself, i.e. whether a certain bundle is not overused by a certain person or in certain corpora type.

Although the studies focus on different aspects, work with different sets of data, under different conditions and in different time, they all acknowledge the fact that MWV are difficult to learn for non-native speakers (Gilquin 2015: 25; Nassaji & Tain, 2010: 402; Yasuda, 2010: 250; Wierszycka, 2013: 82, Yang & Hseih, 2010: 1). They seem to be especially difficult for those that do not have PhV in their mother tongue (Gilquin, 2015: 10). Further, it seems that focusing on comparing native and non-native may be fruitful for language teaching, as the teachers familiar with the findings can draw special attention to the underused but also overused PhV, and thus motivate students to mirror the more native-like usage of these structures. Orientational metaphors seem to be especially useful when introduced in the early stages of learning PhV where it is not an additional burden on language learning.

2.2.3. Avoidance of MWV

Avoidance of MWV is hinted upon by many studies and articles presented so far because they are difficult to learn. Further on, we attempt to investigate the reasons why this

happens with different groups (namely different nationalities) of speakers. Pioneers in this field are Dagut & Laufer (1985), who perceive avoidance as "the reverse side of negative transfer, since learners tend to avoid in L2 those structures that have no parallel in their L1…" (Dagut & Laufer, 1985: 73). It is not that L2 students would be ignorant of such structures, but they rather opt for not using them (ibid.). They test their hypothesis about avoidance in two stages. First, they test native speaker preferences of either PhV or their one-word equivalents by multi-choice gap sentences. The results show that native speakers prefer PhV in fifteen instances out of twenty (ibid.: 74). In second stage, these fifteen preferred PhV are used in testing (i.e. multiple-choice, verb translation test, verb memorizing test) of Hebrew speakers of English to see if their preference matches the one of the native speakers (ibid.: 75). 60 new students are tested for each test. The first, multi-choice test, this time consisting of only the fifteen PhV[12] shows that the Hebrew university students avoid the PhV in 58 % of instances (ibid.). In the translation test (done by partly English majors and partly EFL students), again the verb is left out but is written at the end in Hebrew; in 32 % of the cases English majors opt for the PhV and only in 15 % is this option chosen by EFL students (deemed of lower English proficiency level); overall only 24 % of the sentences are translated by the use of PhV (ibid.: 75/76). A multi choice test thus yields better results, in a sense that more PhV are used when they are one of the options at hand (ibid.: 75). Lastly, the students are tested for memory in that they are given sentences with PhV as well as single words, are instructed to memorize them and in an hour take a test similar to the translational test (ibid.: 76). Although the students are exposed to PhV in context, they only use it in the testing phase in 24 % (as opposed to 41 % single-word equivalents and 35 % faulty or empty slots) (ibid.). The literal PhV are the most frequently chosen in all the tests, followed by completive and lastly, by figurative ones (ibid.: 77). The overall results thus confirm the hypothesis about avoidance. The reason for avoidance Dagut & Laufer (1985) see in the inexistence of V + directional adverb particles in Hebrew which creates the natural tendency for avoiding such structures. They call this avoidance a case of "indirect interference, which results in the learners' inability to comprehend what is being required of them" (ibid.:78).

Liao & Fukuya (2002) replicate the methodology of Dagut & Laufer (1985), this time with Chinese speakers. It was expected that, because Chinese, as well as Hebrew, do not have the underlying structure in their L1, they will avoid using it in English (Liao & Fukuya, 2002: 78). In addition to Dagut & Laufer's (1985) methodology, they take language proficiency into

[12] The students can choose between four options: PhV, single-word equivalent, and two unfitting distractors (Dagut & Laufer, 1985: 75).

consideration for all the three tests (advanced vs intermediate) (ibid.) as well as the impact on performance based on the test type (ibid.: 79). 30 advanced (10 for each test) and 40 (15 for each test) intermediate students are involved in this study. The result show that the division to language levels of students is important, as intermediate students use much fewer PhV than advanced and native speakers (45 % vs 75 % vs 84 % respectively) in the multiple-choice test (ibid.: 89). Overall, the Chinese students opt for the single-word equivalents rather than PhV, however, avoidance (and the "indirect interference" proposed by Dagut & Laufer, 1985) holds true mainly for the intermediate level and it diminishes with increasing level of proficiency (ibid.: 90). Therefore, they suggest that avoidance may not be caused by the absence of V + particles in students' mother tongue, but rather by the proficiency of the speakers; in other words, the more proficient a learner is, the more PhV he/she will probably use, regardless of their L1 (ibid.: 91).

Overall, all the speakers (native, advance Chinese, and intermediate Chinese) favour literal PhV over the figurative ones (ibid.: 93). The performance of advanced and native speakers is found to be closely similar, therefore Liao & Fukuya (2002) claim the advanced students do not avoid any of the two groups of PhV when compared to the native speakers, unlike intermediates that avoid both the groups (ibid.). The preference for the literal over figurative PhV is prominent especially in the translation test (ibid.: 96). The reason may be the semantics of the verb, as the figurative PhV carry the idiomatic meaning, as well as syntactic one, because some verbs connect with different particles and then express different meanings (ibid.).

Houshyar & Talebinezhad (2013) focus on avoidance of PhV by Persian EFL speakers. The conduct is similar to Degut & Laufer (1985) and even more so to Liao & Fukuya (2002) as the aim of the paper is to 1) test PhV use (or lack of it) by (this time) two tests, i.e. multiple-choice and translational test; 2) compare the usage based on proficiency, i.e. this time pre-intermediate vs advanced; 3) focus on the difference of usage between figurative and literal PhV and 4) prove influence of test type on the possible avoidance (ibid.: 238). The same 15 phrasal verbs are used as in Liao & Fukuya (2002), as well as native speakers are used as a control group (ibid.: 240). It is found that all the four non-native groups (2 based on proficiency and 2 divided by the tests) use less PhV than the natives (ibid.: 243). The advanced students doing the multiple-choice test opt for the most PhV when compared to the remaining three, non-native groups (ibid.). The test-type is not found significant for the performance of the pre-intermediate students (ibid.). The results of Houshyar & Talebinezhad (2013) further agree with Liao & Fukuya (2002) in that the proficiency was proven also to play a role, the advanced

students outperformed the pre-intermediate students (ibid.: 245), and in that there is a preference for literal PhV over figurative ones (ibid.: 246). For the last part, they do not find any correlation between higher avoidance resulting from the test-type (ibid.: 247).

Persian EFL learners are also the group studied by Barekat & Banisady (2014) for their avoidance of PhV, and how this pathology has an effect on their written production (Barekat & Banisady, 2014: 346). The first stage is the same as for Liao & Fukuya (2002) and Dagut & Laufer (1985), the students are divided into groups, each group doing one of the three traditional tasks: multiple-matching, translational test, and recall test (Barekat & Banisady, 2014: 346). Overall, results show the EFL students prefer the single-word equivalents over PhV (ibid.: 347). Based on the tests, two groups are formed: group A who scored below the mean score of the tests (avoided more), and group B that scored above (avoided less) (ibid.). After the tests are performed, the two groups are given a writing task: "If I Had a Million Dollars" (ibid.: 346) with the aim of finding out whether the quality of their writings depend on their group affiliation. As the results show, group B scored higher than group A on the writing task. The authors suggest therefore that "there exists a close relationship between phrasal verb avoidance and the participants' writing ability. In fact, phrasal verb avoidance negatively affects the participants' writing performance" (ibid.: 348).

De Cock (2006) also acknowledges the problem of PhV avoidance in production of EFL speakers. She uses the power of ICLE[13] and LINDSEI corpora to provide evidence for her claims. De Cock (2006) proves again what has been also presented by other authors: students who do not have PhV in their mother tongue tend to avoid using them in English (i.e. use fewer PhV than natives and supplement them with single-word equivalents) (ibid.). On the other hand, when speakers do have them in their mother tongue (e.g. Dutch or German) they tend to use them even more than the natives (ibid.). It is not only the avoidance of PhV, however, that the non-native speakers struggle with; the existence of PhV also causes problems in other areas. Stylistic deficiency is one of them, as students are found to use more PhV in formal writings than in informal settings (ibid.), consequently using informal PhV in formal settings (ibid.). This situation probably arises due to the fact that the EFL students have more time to think while writing and rather opt for a piece of language they are not usually comfortable with in speech (ibid.). Further, mistakes are often caused by semantic confusion, i.e. "incomplete understanding of the meaning" (ibid.), namely by confusing PhV with the related single-words, e.g. *Students couldn't put on (wear) a scarf in winter* (ibid.), by choosing the wrong particle, e.g. *They fill up (fill in) many forms* (ibid.), and by choosing the wrong verb, e.g. *Saddam*

[13] International Corpus of Learner English (De Cock, 2006)

Hussein had the power to shut off (turn off) the heat in millions of homes (ibid.). Collocational awareness also seems to be a problem connected to PhV as some seem to be less relevant than others (or not relevant at all) in different contexts, e.g. …*teaching the moral values and preparing them to set up (start) their own family* (ibid.). Students also create their own PhV that do not really exist in the native setting, e.g. *these differences need to be levelled down (ironed out)* (ibid.). Lastly, it is often the case that students use transitive verbs intransitively and the other way around, e.g. *The state should help parents to grow up /* intransitive *(bring up/*transitive*) better generation* (ibid.).

Besides PhV, students sometimes also struggle with PrV. One of the reasons is that the verb is not PrV in their mother tongue and thus they do not treat it as such in English: *I would also like to comment (comment on) the second part of the title* (French speaker). The opposite is also problematic: *And at the same time, he is courting to (courting) a lady* (Spanish) (ibid.). In other cases, PrV is in both languages, however it differs in the preposition used: *It depends of our mental image of the matter* (ibid.). It is not always the mother tongue that causes an issue, sometimes it is the difference in the target language itself, i.e. the case of "interlanguage confusion", e.g. verb creates PrV with various prepositions; deverbal noun comes with its inherent preposition but the verb itself does not come with the preposition (*discuss* vs *discussion about*), verbs often function as single as well as multi-word PrV with a change of meaning (*attend* vs *attend to*), and confusion between *to* being an infinitive marker and a preposition (*She had consented to marry him…* vs S*he had consented to marrying him*) (ibid.). Lastly, style deficiency again plays a role, as EFL students use inappropriate PrV in formal settings: *But the English version of the Treaty talked about (mentioned) land ownership* (ibid.).

The findings reflect the suggestion for practical teaching PhV and Prv, which should be done in "chunks" (ibid.)., i.e. the MWV should not be presented as isolated units but rather in real contexts with a focus on their formal appropriateness, collocation and syntactic place (as is also suggested by Trebits, 2009); if the students do not have MWV in their mother tongue, even more time should be devoted to studying the structures and their position in the English language (ibid.). Although not solely focused on avoidance, De Cock's (2006) article brings to the light the obstacles learners must overcome when mastering MWV which, ultimately, lead to the avoidance and incorrect usage of these structures, proposing ways how to remedy that.

Other studies also focus on the avoidance phenomenon, e.g. You's (1999) study focuses on PhV avoidance of Korean speakers, the avoidance by Egyptian EFL learners is covered in Abdel Salam El-Dakhs (2016), and that of Iranian speakers by Ghabanchi & Goudarzi (2012).

All the researchers discover the tendency to avoid PhV, alluding to the L1-L2 difference being the unifying cause (among other additional ones[14]) of the issues with PhV (Ghabanchi & Goudarzi, 2012: 43; Abdel Salam El-Dakhs, 2016: 132; You, 1999: 153), However, for the reason of space, they could not be discussed any further here.

To summarize, the three research areas provide another insight into the problematic topic of MWV, as well as offer possible ways how to make the issue in question more accessible for students and teachers alike. It can be concluded that:

1) It is important to point out the most frequent MWV and teach them in natural situations and contexts as those are the ones the students will encounter the most.
2) There is a need for a clear knowledge of the structure as well as for a semantic differentiation based on contexts.
3) Avoidance of MWV is mainly caused by indirect interference, arbitrary exposure, and students' lower proficiency in English. It is therefore essential that MWV are taught systematically and in the early stages of students' learning period.

[14] "Semantic complexity (Ghabanchi & Goudarzi, 2012: 43; You, 1999: 151), "passive learning for comprehension and limited language exposure" (Abdel Salam El-Dakhs, 2016:132), and "educational methodology (You, 1999: 151), i.e. rule-simplification strategy" (ibid.: 153) seeing the single word and PhV as mere synonyms in all contexts (ibid.: 151).

3. MATERIAL AND METHOD

3.1. HYPOTHESIS 1

Hypothesis 1 is based on a) our instinctive presumption and b) supported by findings of other authors presented in chapter 2.2.: as students of English overall are found to have difficulties with MWV (Gilquin 2015: 25; Nassaji & Tain, 2010: 402; Yasuda, 2010: 250; Wierszycka, 2013: 82, Yang & Hseih, 2010: 1) which often lead to their avoidance (Dagut & Laufer (1985), Liao & Fukuya (2002), De Cock (2006)), it is presupposed that the studied group of native speakers will use all MWV, i.e. PhV, PrV and PPV more frequently compared to the non-native, Czech speakers studied in this thesis.

3.2. HYPOTHESIS 2

Hypothesis 2 elaborates on the frequency of use/disuse of MWV in that it is expected that non-native speakers will use more PrV over the two remaining classes (PhV and PPV) as PrV are mostly not as opaque in their compositional meaning as PhV and PPV are. One of the reasons for avoidance of PhV may be the students' feeling of inappropriacy of using PhV in quite a formal context, albeit a friendly one because the PhV may be seen as too "colloquial, casual, informal" (Wierszycka, 2013: 89); the preference for rather non-idiomatic structures (i.e. seeing phrasal verbs as more idiomatic than prepositional verbs) is therefore expected.

3.3. HYPOTHESIS 3

Hypothesis 3 is based on the phenomenon of the so-called "phrasal teddy-bears" (Ellis, 2012: 17) or "phraseological teddy bears" (Hasselgård, 2019: 340) i.e. learners of a language have a tendency to overly rely on very frequent chunks of a target language, to an extent that this reliance surpasses a normal usage of these chunks by native speakers. In addition, they also overgeneralize them and thus use them in context that would not be typical for native speakers (ibid.). As the non-native speakers in this study range from B2 - C2 language level (and thus most of them are still learners), this may also be the case with the usage of MWV. It is therefore hypothesised that the non-native, Czech speakers will favour certain MWV (i.e. PhV, PrV, or PrV) that are far less frequent in the native data i.e. the non-native speakers are expected to have their favoured teddy-bear MWV which they will use significantly more than the native speakers. If this happens to be the case with certain MWV in this study, we also hypothesis that the non-native speakers will use the MWV in more contexts than it is in norm with the native speakers.

3.4. MATERIAL

The data for analysis come from two interconnected sources: LINDSEI_CZ, i.e. *Louvain International Database of Spoken English Interlanguage* restricted to the Czech speakers only and LINDSEI's referential corpus LOCNEC, i.e. *Louvain Corpus of Native English Conversation.*

LINDSEI is a sister counterpart to ICLE (*International Corpus of Learner English*); unlike ICLE, which is a corpus compiled of written works of university students, LINDSEI is a corpus of spoken interactions. Although originating in Louvain-la-Neuve in Belgium, the corpus consists of spoken interviews of university students from eleven countries (DeCock, Guilquin, Granger, 2010: 3); in other words, there are eleven sub-corpora of LINDSEI that differ in mother tongues of the students-speakers.[15]

Two people participate in each of the interviews: one interviewer (a teacher or a tutor) and an interviewee (a university student). The talk is divided into three parts: a warm-up activity on a given topic (students choose from three options[16]), an informal conversation targeted on student's life, hobbies and interests, and a picture description[17] (ibid.: 8-9). The interviews cannot be seen as completely natural, strictly speaking, as the students are informed about the task and its structure beforehand and there is an aim behind the interaction. However, the warm-up activity as well as the informal conversation come close to being seen as data from the typical learner corpora (ibid.: 6), with the picture description being more controlled and constrained by requirements (ibid.).

All the speakers participating in the LINDSEI corpus must fit in required criteria for easier interpretation of results, i.e. university undergraduates, EFL rather than ESL learners with advanced proficiency level[18] (ibid.: 7). Other variables are also available and can be taken into account for analysis, e.g. gender, mother tongue, or knowledge of other foreign language (ibid.: 11-12). All the interviews (usually around 50 per sub-corpus) are held in an informal

[15] There were eleven corpora created when the booklet was published in 2010. However, the corpus has expanded to incorporate other languages over the years, Czech among them.
[16] A) An experience you have had which has taught you an important lesson. You should describe the experience and say what you have learnt from it.
B) A country you have visited which has impressed you. Describe your visit and say why you found the country particularly impressive.
C) A film/play you have seen which you thought particularly good/bad. Describe the film/play and say why you thought it was good/bad.
[17] Students retell a story in their own words, based on four pictures depicting a woman unhappy with her realistic painting.
[18] After the compilation and subsequent analysis, it was found that the level ranges between higher intermediate to advanced.

tone (ibid.: 7) and are planned to be around 15 minutes long (ibid.: 8), yielding around 2000 words each (ibid.).

The spoken records are transcribed based on a strict blueprint; the final outcome of the transcription is presented in Figure 2.

<h nt="CZ" nr="CZ002">
<S>
<A> hello <name of the interviewee> right what are you <overlap /> going to speak to us about </A>
<B> <overlap /> hi (em) I've chosen the topic three . (erm) I should probably . read out . <overlap /> the[i:] instruction right . yeah </B>
<A> <overlap /> yes read out the topic and </A>
<B> (em) . (er) I will be talking about a film or a play that I've seen . which I thought was particularly good or bad . and then describe it </B>
<A> great . off we go </A>
<B> (em) so (em) . last week I went with a with a couple of my colleagues (eh) to <foreign> Čino= Činoherní klub </foreign> and (em) we saw (eh) Glengarry Glen Ross . which is a brilliant play by David Mamet . my David Mamet that I wrote my BA thesis on </B>
<A> wow </A>

Figure 2 – illustration of LINDSEI transcript (CZ002)

<A> and </A> mark the beginnings and ends of the interviewer's tiers, whereas <B> and </B> mark the interviewee's ones (ibid.: 17-18). For the final analysis, only the <B> tiers are used.

The second corpus used, LOCNEC, has been compiled as a "comparable native speaker corpus" (ibid.: 65). It is transcribed and structured the same way LINDSEI is, i.e. warm-up activity, free conversation and picture description (ibid.). The participants are also university students (from Lancaster University, UK) (ibid.).

3.5. METHOD OF ANALYSIS

The core interest in the Analysis are multi-word verbs (MWV), i.e. phrasal verbs (PhV), prepositional verbs (PrV) and phrasal-prepositional verbs (PPV) and their similarity and difference of use by native (British) and non-native (Czech) speakers of English. Before the analysis of MWV could have been realized, several steps had to be taken to enable it. Transcripts of 50 Czech speakers from LINDSEI_CZ and of 39 native speakers from LOCNEC were obtained. Then, each transcript was divided into two separate documents, one with only tiers <B> and one with exclusively tiers <A>. This has been done in order to search for and compare only the use of MWV by students in the final analysis, without an interference of the interviewers' turns. The data, i.e. the transcripts with either only <A> or <B>, or both <A> and <B> from LINDSEI_CZ and from LOCNEC were tagged for parts of speech in the Sketch Engine program. Therefore, six separate sub-corpora of both LINDSEI_CZ and LOCNEC are created, i.e. LINDSEI_CZA, LINDSEI_CZB, LINDSEI_CZAB, LOCNECA, LOCNECB, and LOCNECAB. This has been done for several reasons: the corpora with <B> tiers only are analysed first, solely focusing on all the potential MWV used by the students; then corpora with

<A> tiers only and both tiers present were analysed second, giving MWV that the student potentially only repeated after the interviewer. These repeated MWV are therefore disregarded from the final analysis, as the students could only echo the teacher.[19] The complete corpora, i.e. LINDSEI_CZAB and LOCNECAB also provide the total number of words that appear in the data when compared to the LINDSEI_CZB/LOCNECB that are used as references in the final analysis.

After the sub-corpora are compiled in the Sketch Engine program, a query is used to search for the desired MWV. First, we look for a verb (V.*) that is followed by either a preposition (IN), an adverb (RB.?) or by a particle (RP) with a possibility of there being maximum of three repetitions of either a determiner (DT), an adjective (J.*), a noun (N.*), or a pronoun (PP.*) or their combinations in between the verb and the preposition/particle; thus the query is structured as follows:

[tag="V.*"][tag="N.*"|tag="PP.*"|tag="DT"|tag="J.*"]{0,3}[tag="RP"|tag="RB.?"|tag="IN"]

Through this query, it is possible to search for potential phrasal and prepositional verbs that may or may not have an element between their verb and particle. The second step is to search for phrasal-prepositional verbs through a separate query[20]

[tag="V.*"][tag="N.*"|tag="PP.?"|tag="DT"|tag="J.*"]{0,3}[tag="RP"|tag="RB.?"][tag="IN"]

These two separate queries are opted for, in order to make the manual searching less overwhelming and thus easier.

The queries mentioned above search for the desired combinations of words, yielding 8 559 potential MWV in LINDSEI_CZB and 10 354 in LOCNECB. It is clear at a first glance that not all the resulting tokens can be considerate as potential MWV, e.g. *according yeah, amazed by, were just,* etc. (LINDSEI_CZB), therefore a manual selection of truly potential MWV had to be made. This manual check decreased the number of potential MWV to 866 and 1276 tokens respectively. After this manual check the potential MWV are consulted with several sources, i.e. dictionaries[21], Biber et al (1999), Quirk et al. (1985) and a dictionary of phrasal verbs, i.e. *Longman Phrasal Verbs Dictionary* (2000). Their definition of a phrasal verb is "a verb that consists of two or three words. Most phrasal verbs consist of two words [of which] the second word is a particle. The particle is either an adverb or a preposition [e.g. *get up, deal with*]. There are also some three-word phrasal verbs, for example *catch up with*… An important

[19] These MWV were disregarded when the student had repeated them in their following sentence
[20] The only difference is that the second query searches for a particle or an adverb which is followed by a preposition
[21] i.e. Online Cambridge Dictionary (2020), Macmillan Dictionary (2020), and Collins Dictionary (2020)

feature of phrasal verbs is that they are typically idiomatic" (Summers, 2000: xi). Clearly, this covers all that we distinguish here as phrasal, prepositional and phrasal prepositional verbs. This dictionary has proven useful after the manual selection as it gave a guidance to ambiguous cases of MWV and mainly provided the various semantic meanings with certain polysemous MWV. However, this is not to say that only MWV that appear in the dictionary are considered further, as some MWV are in the final analysis but are not present in the dictionary; however, they are present in grammar books e.g. *think about, ask for* etc. When a potential MWV is not incorporated in the dictionary but could be found in the grammar books, the interpretation of the grammar books is mostly favoured.[22] Lastly, before the MWV were classified into their designated categories, all the MWV had been checked again in the respective corpora to ensure maximum correctness of token frequency. The query used for this:

[lemma="X"][tag="N.*"|tag="PP.?"|tag="DT"|tag="J.*"]{0,3}[lemma="Y"]

searches for a given specific verb (marked here as X), e.g. *give* and a given specific preposition/adverb/particle (marked here as Y), e.g. *up* with there being a possibility of a noun, pronoun, determiner, or adjective in between the two lemmas.

The process of decreasing the number of potential MWV with the data concerning the overall and <B> line word-count is summarized in Table 3:

Corpus	**LINDSEI**	**LOCNEC**
Total number of words, <A> and <B>	127 827 tokens	143 785 tokens
Number of words, <B> tier only	98 814 words / tokens	106 975 words / tokens
Results of the queries (<B> tier only) Potential PhV, PrV and PPV	8 559 tokens	10 354 tokens
Number of MWV after manual selection	207 types / 866 tokens	292 types / 1276 tokens
Number of MWV after the dictionary consultation and another manual check	170 types / 404 tokens	261 types / 959 tokens
Number of MWV for final analysis	170 types / 622 tokens	245 types / 1153 tokens

Table 3 –potential and final MWV in the corpora

[22] Another reason for siding with the grammar books consulted in the theory is the audience for which the grammar books and dictionaries are written, i.e. grammar books for linguistic purposes whereas dictionaries for educational purposes of general public.

The difference between the number after the dictionary consultation and the final number of MWV is influenced mainly by consultation with grammar books, mainly Biber et al. (1999).

When assuming the significance value, for $p<0.05$, the final number of type and tokens of MWV in the LOCNECB corpus is significantly higher than the number of MWV tokens in LINDSEI_CZB. Therefore, the difference is statistically significant between the corpora.

After reaching the final number of MWV for analysis, these MWV were further classified into their designated classes: phrasal verbs, prepositional verbs and phrasal-prepositional verbs. This has been done first based on the particle used, i.e. whether it is a preposition (denoting PrV) or spatial and locative adverb (denoting PhV) or functioning as both. The classification of particles is done based on Quirk et al. (1985: 1151), which has been described in chapter 2.1. by Figure 1. The representation of given MWV in the final analysis is based on their form, not their separate meanings, e.g. *go on* denotes 'happening', 'encouragement' or 'continuation', however, it has been analysed based on form as one MWV (i.e. PhV).

In cases when the particle works both as a preposition and an adverb, i.e. it is a prepositional adverb, a class B particle (cf. Figure 1), the scale of idiomaticity has been tested: more idiomatic MWV with type B particle were grouped with PhV, whereas more transparent MWV were grouped with PrV. Besides idiomaticity, we relied on the syntactic tests distinguishing PhV from PrV. These tests have been presented in chapter 2.1.4. (e.g. passivization, the position of NP, placement of adverbial particle etc.). In order to distinguish between FC and true MWV these tests were also applied, especially the test of *wh*-questions indicating either PrV or FC. In the case of *come from* (and the like), the question is formed with *where*, i.e. *where do you come from?* Thus, based on the tests, *come from* should be seen as FC. However, most dictionaries consulted see it as MWV. Therefore, in such dubious cases when the tests did not coincide with the dictionaries or grammars, we opted for siding with the academic sources.

All the steps taken in order to acquire the final list of MWV from the raw data of LINDSEI_CZ and LOCNEC are summarized in Table 4.

Method of extracting MWV	
Step 1	the transcripts divided based on <B> and <A> tiers for both the corpora
Step 2	Sketch Engine tagged all the transcripts for parts-of-speech
Step 3	the queries searched for the desired combinations of words
Step 4	manual selection of truly potential MWV from the results of queries
Step 5	the potential MWV consulted with a dictionary of phrasal verbs + grammar books
Step 6	Check of token numbers through the third query using lemmas
Step 7	Final number of MWV
Step 8	Categorization of MWV into their three designated types

Table 4 – method of extracting MWV

Several aspects have to be considered regarding the analysis itself and its final data before the results of especially quantitative analysis are discussed. First, the data themselves bring an issue to the analysis, as they are transcripts of speech and as such may be burdened by the quality of the recording. This issue may cause the transcriber to have difficulties hearing the correct word. Also, the transcriber him/herself may have misunderstood the words or simply may not have known them. Therefore, the data are to some extent, influenced by the transcriber as well.

The second issue also concerns the data, especially its amount and the need for a heavy manual selection. As the numbers show, the queries used yielded 8559 tokens in LINDSEI_CZB and 10 354 tokens in LOCNECB, however, after manual check, a lot of the combination had to be disregarded from the analysis, and the number of tokens decreased quite drastically, to 866 and 1276 tokens respectively, even more so after the following dictionary consultation and further testing, to the final 622 tokens (i.e. 170 types) in LINDSEI_CZB and 1153 tokens (i.e. 245 types) in LOCNECB. Although, as the overall claims in the analysis are conclusive in that there is a clear difference between the two sets of data we allow for a human error in the qualification of the data.

4. ANALYSIS

The aim of the Analysis is to discuss the findings regarding MWV in the two corpora LINDSEI_CZ and LOCNEC. It first discusses the use and frequency of MWV, i.e. phrasal, prepositional and phrasal-prepositional verbs in the LINDSEI_CZB corpus, followed by findings in the corpus LOCNECB. Lastly, the findings from both corpora are compared in order to show the difference of usage overall, as well as of the given categories.

4.1. MWV IN LINDSEI

Overall, there are 170 MWV types, which accounts for 622 MWV tokens in the LINDSEI_CZB sub-corpora (which itself consists of 98 814 tokens total). All the researched subclasses of MWV are present in the data; their frequency is summarized in Table 5 below.

MWV in LINDSEI_CZB				
	Types	Representation in %	Tokens	Representation in %
TOTAL	170	100 %	622	100 %
PrV	80	47,1 %	398	64 %
PhV	77	45,3 %	191	30,7 %
PPV	13	7,6 %	33	5,3 %

Table 5 – MWV distribution in LINDSEI_CZB

When types are counted PrV are the most frequent: 80 different types of PrV are found in the LINDSEI_CZB corpus. There are 77 cases of PhV and 13 cases of PPV. As has been expected, PPV are the least frequent of MWV as there is overall less variety in comparison with other MWV and their structure is quite complex and idiomatic in nature, requiring both a particle/adverb followed by a preposition. The number of PrV types is comparable to that of PhV; there is a three-type difference between the two MWV. What is more telling, though, is the number of tokens in the corpus. PhV are represented by 191 tokens in the corpus, whereas there are 398 tokens of PrV, indicating the dominance of PrV, which is statistically significant at $p<0.0001$. This is also supported by the type/token ratio, which has been counted 0.20 for PrV and 0.40 for PhV. Therefore, the students use around as many PhV as PrV types, however, they use PhV less frequently token-wise. This may be caused by the opaque nature of PhV, and difficulty with retrieving them from their mental lexicon, as they are probably less deep-rooted and readily at hand as PrV are.

First, 50 most frequent PrV found in the LINDSEI_CZB are summarized in Table 6[23]. Overall, there are 80 types of PrV which amounts to 398 tokens.

[23] The complete lists of all MWV found in the corpus are in the Appendix

50 Prepositional verbs – LINDSEI_CZB			
MWV type	**raw frequency**	**MWV type**	**raw frequency**
Talk about	40	Feel like	4
Think about	38	**Give to**	4
Look at	36	Go with	4
Look like	21	**Know about**	4
Go for	20	**Remind of**	4
Listen to	14	**Base on**	3
Talk to	12	Care for	3
Focus on	11	Concentrate on	3
Say to	10	Consist of	3
Wait for	9	Divide into	3
Work on	9	Get over	3
Come from	8	Happen to	3
Get into	8	Laugh at	3
Turn into	8	**Turn to**	3
Be into	7	Come across	2
Think of	7	Engage with	2
Deal with	6	Get through	2
Apply for	5	Hear of	2
Go through	5	Look after	2
Point at	5	Play with	2
Say about	5	Search for	2
Speak to	5	Spy on	2
Connect with	4	**Switch into**	2
Depend on	4	**Accuse of**	1
Do about	4	**Ask for**	1
TOTAL	301	**TOTAL**	67

Table 6 – fifty most frequent prepositional verbs in LINDSEI_CZB

Those PrV that are in bold tend to fall into TypeII PrV, i.e. they require Od as well as a prepositional object. There are 26 such types that enter the required frame, represented in Table 6 by, e.g. *base sth on sth, ask sb for sth*, or *say sth to sb*. The remaining cases of PrV rather conform to TypeI PrV, requiring only the prepositional object. However, certain PrV seem to appear in both frames; the use depends on the structure of a given sentence, omission or context. Therefore, when the PrV allows for both the Types, i.e. TypeI and TypeII, it is counted as TypeII for clarity, although we acknowledge the possibility of the Type I frame being also valid in many cases. For example, *turn to* is found to function as both Type I and TypeII; other such PrV are e.g. *turn into, begin with, say to* or *write to*

(1) *As I came back from the united states I'd felt […] I need to get . back to European identity [so] I decided to* ***turn*** *it* ***to*** *British English* (LINDSEI_CZB)

(2) *Phonology seminars […] helped a lot so I* ***turned to*** *British English* (LINDSEI_CZB)

(3) *They* ***say*** *goodbye* ***to*** *each other at the (er) at the train station* (LINDSEI_CZB)

The most frequent PrV in the corpus is *talk about* (40 occurrences), followed by *think about* (38 occurrences) *look at* (36 occurrences)), and *look like* (21 occurrences).

(4) *...and they start **talking about** this but then of course the conversation becomes about anything but the children* (LINDSEI_CZB)

(5) *...can I take some time to **look at** it or do I have to speak right away* (LINDSEI_CZB)

48 PrV appear more than once in the corpus (they are all presented in Table 6), e.g. *deal with* (6 occurrences; sentence 6), *know about* (4 occurrences) and *come across* (2 occurrences). Therefore, 60 % of PrV is represented by at least two tokens.

(6) *Mamet was one of the playwrights that we **dealt with*** (LINDSEI_CZB)

In addition, some of the numbers had to be reduced for the reasons of students' repetition after the interviewer (e.g. *talk about;* sentence 7*, talk to*) and also because the instances were not denoting PrV but rather FC. In fact, *talk about*, which is the most frequent PrV, appears in the corpus LINDSEI_CZAB 102 times, with 66 instances in the LINDSEI_CZB; however only 40 cases that are not repeated are considered for the quantitative analysis

(7) *<A> what have decided to **talk about** </A>*
*<B>I decided to **talk about** a country . which has impressed me </B>* (LINDSEI_CZAB)

Additionally, some instances of V + preposition are disregarded because of their clear usage as FC, e.g. *go through* (sentence 8) *look into* (sentence 9, first part) or *get into.* These are, however, counted for in the analysis when they are used in their more idiomatic sense (e.g. *look into ~study*). Some potential PrV appear only in FC in the corpus and thus do not appear at all in the analysis, e.g. *enter into.*

(8) *...he just **goes through** the: . through the[i:] old Prague* (LINDSEI_CZB)

(9) *When we **look into** (er) the mirror we are not usually very satisfied* vs *I really like to **look into** the depth (er) of the language* (LINDSEI_CZB)

(10) *[Pantheon]'s huge it's massive and when you when you really **enter into** it it's [] it was really and amazing experience* (LINDSEI_CZB)

The results also show that all the semantic categories, i.e. communication (e.g. *ask for*) activity (e.g. *look for*), mentality (e.g. *think of*), causativeness (e.g. *come from*), simple occurrence (e.g. *look like*), existence/relationship (e.g. *base on*), and aspect (e.g. *begin with*) (Biber et al.,1999: 408) are present in the corpora. When the token representation is considered, the most frequent categories are activity, mentality and communication which is in agreement

with the findings by Biber et al. (ibid.: 419). The results are also given by the design of the corpus where the main objective is communicating experience with certain activities and the feelings resulting from them.

77 PhV types are found in the LINDSEI_CZB corpus, with 50 being presented in Table 7. The 77 types amount to 191 tokens.

50 Phrasal verbs – LINDSEI_CZB			
MWV type	**raw frequency**	**MWV type**	**raw frequency**
Find out	18	Go down	2
Come back	17	Go up	2
Go on	11	**Look around**	2
Go back	9	Meet up	2
Turn out	9	Move out	2
Go out	7	**Put on**	2
Pick up	7	**Set up**	2
Give up	4	Start off	2
Go over	4	**Take up**	2
Sit down	4	**Throw out**	2
Work out	4	**Build up**	1
Carry on	3	**Burn down**	1
Get back	3	**Call up**	1
Grow up	3	Catch up	1
Hang out	3	**Check over**	1
Help out	3	Come on	1
Show off	3	Come over	1
Stand up	3	Die out	1
Bring up	2	End up	1
Come up	2	Get up	1
Dress up	2	Go away	1
Drop out	2	**Hand in**	1
Figure out	2	**Keep up**	1
Fill in	2	Lie down	1
Get out	2	**Light up**	1
TOTAL	129	**TOTAL**	35

Table 7 – fifty most frequent phrasal verbs in LINDSEI_CZB

PhV TypeI (intransitive PhV) and PhV TypeII (transitive PhV, requiring Od) are both represented in the corpus. There are 36 types that belong to PhV Type I, e.g. *go on, hang out* and *look back*. The matter is more complicated with the remaining 41 PhV that should belong to PhV Type II (transitive PhV, marked in bold), e.g. *pick up, build up* and *bring up*; however, some of them, e.g. *work out* or *get out* are used both transitively and intransitively. It is therefore important to note that with some PhV definite boundaries between the two types cannot be drawn.

(11) *I like **hanging out** with friends* (LINDSEI_CZB)

(12) *...different methods (er) of (er)* ***bringing*** *the children* ***up*** *so that was that was the main problem* (LINDSEI_CZB)

(13) *...that's the plan hopefully it's going to* ***work out*** (*~function*; intransitive) vs *…but sometimes you just got the idea that you could* ***work*** *something* ***out*** *differently* (*~solve/manage*; transitive) (LINDSEI_CZB)

Work out is not the only PhV that has more than one sense (as exemplified by sentence 13). The PhV *go on* appears in the corpus in 22 cases total, however, only 11 of those are classified as PhV, the remaining 11 are disregarded for their complete transparency and non-idiomatic meaning. In addition, the 11 PhV tokens share more than one idiomatic meaning:

(14) *As the play . (erm)* ***goes on*** *(eh) you see them . going at each other's throats* (LINDSEI_CZB)

(15) *There is like nothing* ***going on*** *no actional scenes nothing* (LINDSEI_CZB)

(16) *Studying literature opens your mind [] I think it was great to have the= these three years here . and now I think I can .* ***go on*** (LINDSEI_CZB)

Go on seems to have three semantic meanings in the corpus. *Go on* in sentence 14 denotes a continuation of the play, a progression of the play in a linear direction. There are three instances of this meaning in the corpus. In sentence 16 a kind of progression is also present; however, it rather points to finishing with one thing and moving on to another one; this meaning is expressed by *go on* only once. In sentence 15 *go on* denotes the action of happening; the PhV *go on* and *happen* are interchangeable in this sense; there are seven cases of *go on* denoting happening overall.

The most frequent PhV used by the Czech students is *find out* (having the one-word synonym in *discover* or *realize*) which appears 18 times in the corpus, followed by *come back* (17 occurrences), *go on* (11 occurrences), *go back* (9 occurrences) and *turn out* (9 occurrences).[24]

(17) *I actually* ***found out*** *that . (er) I am able to certain level to understand* (LINDSEI_CZB)

(18) *I know that my accent when I* ***came back*** *it was <starts laughing> definitely better* (LINDSEI_CZB)

[24] With *come back* and *go back* the meaning of the whole can be, to an extent, deduced from the compositional parts and thus e.g. Biber et al (1999) consider them to be FC. However, both *come back* and *go back* have their one-word synonym (i.e. *~return*), the particle *back* has a figurative sense, not denoting 'away', and all the other sources consulted, i.e. the dictionaries, list both of these composites as PhV. For these reasons, these two PhV were included in the analysis for both LINDSEI_CZB and LOCNECB corpora.

35 PhV out of the 77 appear more than once in the corpus, e.g. *figure out* (2 appearances), *go out* (7 occurrences), *carry on* (3 occurrences), and *start off* (2 occurrences)

(19) *...plus I don't have the finances to . to **carry on** here in the[i:] UK* (LINDSEI_CZB)

(20) *...so he starts **going out** with this young girl you know they become a couple* (LINDSEI_CZB)

Go out is also one of the V + particle that had to be analysed in more detail, as only the more figurative meaning of *go out*, i.e. *~date, enjoy yourself* is considered here. For this reason, the total number of tokens has been reduced from eight to seven.

There is also a handful of PhV that can be further classified not only as PhV but intensified PhV, meaning that the particle in these cases seems to add mainly the value of closure or exaggeration and can be often omitted from the sentence, e.g. *end up, print out, open up.*

(21) *...he becomes intrigued by who this . mystery young lady is and then they actually **end up** together* (LINDSEI_CZB)

Overall, PhV found in the LINDSEI_CZB can be assigned to the semantic categories of communication (*point out*), activity (*come over; take up*), mentality (*give up*), aspect (*start off*), and occurrence (*grow up*).

All the 13 types (i.e. 33 tokens) of PPV found in the corpus are presented in Table 8.

Phrasal-prepositional verbs; LINDSEI_CZB	
MWV	**raw frequency**
Look forward to	12
Go back to	5
Come up with	3
Get away from	2
Get back to	2
Get on with	2
Catch up with	1
Come back to	1
End up with	1
Get around to	1
Hold on to	1
Make up for	1
Try out for	1
TOTAL	33

Table 8 – all phrasal-prepositional verbs in LINDSEI_CZB

The most frequent phrasal-prepositional verb is *look forward to* occurring 12 times in the data, always as a marker of pleasant anticipation[25]:

(22) *I was really **looking forward to** period of time after mi finals* (LINDSEI_CZB)

(23) *I'm . usually **looking forward to** going home . for a few days* (LINDSEI_CZB)

Other PPV that appear more than once in the corpus are *go back to* (5 occurrences) *come up with* (3 occurrences), *get on with, get back to* and *get away from* (2 occurrences). The rest of the PPV appear only once in the data:

(24) *...at the BA it was quite difficult to **catch up with** the with the reading schedule* (LINDSEI_CZB)

Interestingly, PPV *put up with* does not appear in the data, although it is taken as a classic example in the grammar books. Overall, activity (e.g. *make up for*), mentality (e.g. *come up with*) and causative (e.g. *end up with*) semantic categories are those represented in the data.

[25] The PPV was completed by the *-ing* V only three times in the corpus, the remaining nine cases are complemented by NP (one time) or by *it*.

4.2. MWV IN LOCNEC

Overall, there are 245 MWV types, which accounts for 1153 MWV tokens in the LOCNECB sub-corpora (which itself consists of 106 975 tokens total). All the desired subclasses of MWV are present in the data, the frequency of which is summarized in Table 9 below.

MWV in LOCNECB				
	Types	Representation in %	Tokens	Representation in %
TOTAL	245	100 %	1153	100 %
PrV	73	29,8 %	426	36,9 %
PhV	155	63,3 %	681	59,1 %
PPV	17	6,9 %	46	4 %

Table 9 – frequency of MWV in LOCNECB

In terms of both types and tokens, the most frequent MWV category is PhV, with 155 types found in the corpus which correspond to a total of 681 tokens. The PhV subclass is two times more frequent (in relation to the number of types) than the class of PrV, with 73 types and 426 tokens. The least frequent MWV is PPV with 17 types present in the corpus, which account for 46 tokens. The difference between the number of PrV and PhV types is quite astounding, even though PhV were expected to be more frequent than PrV; however, not by that many types and tokens. The type/token ratio (i.e. 0,17 for PrV and 0,23 for PhV) shows that although there are fewer types of PrV than PhV, the types are represented, when generalized, by more tokens per type. PhV, on the other hand, have twice as many types, however, the ratio indicates fewer tokens per type. Indeed, more than half of PhV types (i.e. 87) have only one or two token representation. This is probably due to the fact that PhV is a more open group for forming new and unorthodox combinations in real time, whereas PrV form rather a more closed category that cannot be expanded that readily. This also points to the fact that PhV are an integral part of the native speakers' casual speech.

Table 10 gives the 50 most frequent PrV found in the LOCNECB corpus.

50 Prepositional verbs – LOCNECB			
MWV type	**raw frequency**	**MWV type**	**raw frequency**
Look at	55	Take over	5
Think about	29	Feel like	4
Go into	27	Hear of	4
Look like	22	Look through	4
Talk about	20	Appeal to	3
Think of	17	Care for	3
Come from	16	Concentrate on	3
Apply for	13	Deal with	3
Say to	12	Look after	3
Write to	12	Relate to	3
Go for	11	**Say about**	3
Look for	10	Aim at	2
Speak to	10	Be into	2
Worry about	10	Come across	2
Ask for	9	Cope with	2
Depend on	9	Focus on	2
Get on	8	Fuss over	2
Do about	7	Go with	2
Give to	7	Hear from	2
Talk to	7	Live for	2
Go through	6	Live on	2
Happen to	6	**Make sth into sth**	2
Look into	6	Rely on	2
Get into	5	Abide by	1
Know about	5	**Base on**	1
TOTAL	339	**TOTAL**	64

Table 10 – fifty most frequent prepositional verbs in LOCNECB

Overall, 73 PrV types are found in LOCNECB. Both PrV TypeI (requiring only prepositional object) and PrV TypeII (requiring by default both Od and prepositional object; marked as bold) appear in the corpus. There are 14 types of PrV Type II[26], e.g. *ask for* and *make into*; the remaining PrV Type I favour the frame with just a prepositional object, e.g. *cope with* or *depend on*

(25) *Tom Cruise* as <name of a character> *decided that you know he'd* ***make*** *him* ***into*** *a vampire…* (LOCNECB)

(26) *I couldn't* ***cope with*** *that no that my brain too much you know* (LOCNECB)

[26] If the PrV allows for it, these PrV TypeII were sometimes used in the sentence without the required Od. The PrV were considered as Type II mainly when there is a possibility of them being in the frame with Od and a prepositional object

48 PrV out of 73 are found to appear at least two times in LOCNECB (all are presented in Table 10). The most frequent PrV is *look at*, with 55 appearances in the corpus, followed by *go into* (27 appearances) and *think about* (29 appearances).

(27) *It was quite interesting and the more I think we* ***looked at*** *the[i:] adverts we actually learnt something…* (LOCNECB)

Go into appears in total of 48 cases in LOCNECB, however, only 27 conform to the idiomatic meaning of 'starting in a new field of interest/job'; in the remaining 21 cases *go into* was used in FC, as 'go into a place'

(28) *There's two ways of* ***going into*** *teaching* (LOCNECB) vs *He says grandpa Joe to* ***go into*** *this chocolate factory and one by one the kids do something wrong* (LOCNECB) (not counted in the final analysis)

Think about, as well as *talk about* are the only PrV that are reduced in number because of students' direct repetition after the interviewer. However, this happened only once for both *think about* and *talk about*. This is exemplified by sentence (29)

(29) *<A> it would be about time to* ***think about*** *it <\A>*
<B> no I ***think about*** *it sooner or later . I'll get round to graduating first<\B>* (LOCNECAB)

Some instances of V + preposition are disregarded because of their usage as FC, e.g. *live on*, (sentence 30), *go for* (sentence 31), however in other instances they appear in the analysis as PrV with their more idiomatic meaning (e.g. *live on ~ survive because of something*).

(30) *No I* ***lived on*** *campus for one year* (FC) vs *I have potentially three thousand three hundred a year to* ***live on*** (PrV) (LOCNECB)

(31) *I mean it's been* ***going for*** *thirty years* (LOCNECB)

The results also show that all the semantic categories, i.e. communication (e.g. *say to*) activity (e.g. *deal with*), mentality (e.g. *know about*), causativeness (e.g. *come from*), simple occurrence (e.g. *happen to*), existence/relationship (e.g. *rely on*), and aspect (e.g. *begin with*) (Biber et al.,1999: 408) are all present in the corpora. If the token representation is considered, the most frequent categories are activity, mentality and communication which coincides with the findings in LINDSEI_CZB as well as with Biber et al. (1999: 419).

Overall, there are 155 PhV types in the LOCNECB corpus. 50 PhV that appear in the corpus most numerously are presented in Table 11 below.

50 Phrasal verbs – LOCNECB			
MWV type	**raw frequency**	**MWV type**	**raw frequency**
Go back	50	**Show off**	8
Come back	42	Turn out	8
Go out	27	Get together	7
Go on	26	**Set up**	7
Work out	18	**Start off**	7
Get out	17	**Take off**	7
Get up	16	Come on	6
End up	14	Move out	6
Get back	14	**Put together**	6
Come up	13	**Sort out**	6
Come in	12	Come over	5
Go away	12	Come round	5
Sit down	12	Fit in	5
Go over	11	**Beat up**	4
Get away	9	**Bring in**	4
Pick up	9	**Build up**	4
Stand up	9	Come through	4
Take out	9	**Cut off**	4
Wake up	9	Grow up	4
Carry on	8	Look around	4
Find out	8	**Make up**	4
Get in	8	**Put off**	4
Go off	8	Stop off	4
Move away	8	**Take on**	4
Put on	8	Come along	3
TOTAL	377	TOTAL	130

Table 11 – fifty most frequent phrasal verbs in LOCNECB

In 103 cases a given PhV appears more than once in the corpus. Those occurring only once but are worth mentioning (because of their rareness or inventiveness) are e.g. *turf out, queue up, pop off, belt out* or *bolster up*

(32) *They get drunk and then when they're **turfed out** of the pubs they just go around* (LOCNECB)

The most frequent PhV in the corpus is *go back* with 50 occurrences, followed by *come back* (42 occurrences), and *go out* (27 occurrences).

(33) *I spent [money] in South America **came back** and worked as a baker* (LOCNECB)

Both PhV Type I and PhV Type II (words in bold) are present in the results, with 77 inclining to the PhV Type II, e.g. *put on* (8 occurrences), *start off* (7 occurrences), *take on* (4

occurrences) and the remaining 78 rather favour the intransitive frame, which is the frame of PhV Type I, e.g. *come up* (13 occurrences), *carry on* (8 occurrences), *move away* (7 occurrences)[27]

(34) *We all **put on** like the tight lycra stuff* (LOCNECB)

(35) *... I **moved away** [] and went to do (er) forestry* (LOCNECB)

As the theoretical sources suggest, *put on* is found to have more than one semantic meaning in the corpus. Sentence 36 illustrates the meaning of 'wear'; other meanings present in LOCNECB are 'start to play', i.e. *put Braveheart on*; 'do an activity', i.e. *they didn't put on dance*; 'add' in *put on weight*[28] (LOCNECB). *Put on* is not the only PhV that have more than one semantic meaning[29]; other such PhV are, for example, *take off*, *take out*, *go off (leave* vs *make noise)*, *make up*. *Make up* denotes either 'fabulize' or 'create/prepare'

(36) *He just **makes** the things **up** as he goes along* vs *we had to work in a group and **make up** our own . half an hour play* (LOCNECB)

(37) *You know don't just **go off** for any odd <?> reason* vs *...then suddenly the fire alarms **went off**...* (LOCNECB).

Take off denotes either 'become successful' or 'remove' and *take out* gives the meaning of 'remove' 'hang out with' or 'get money' in LOCNECB. What is even more interesting, however, is that besides having these various meanings on their own, *take off* and *take out* also share a particular meaning denoting 'have a break from work/life/studies'. Based on the (admittedly limited) examples, it seems we *take out* a whole year, whereas *take off* also works with shorter amount of time, i.e. day, as well as a year.

(38) *I might **take** a year **out** in France* (LOCNECB)

(39) *They sort of you know they **take** a day **off** work to go to the bank* (LOCNECB)

As has been expected, there are some PhV that can be categorized as intensifying, i.e. the particle intensifies the meaning of the verb and to an extent, adds a characteristic of closure (indicating that an action has ceased to happen) or emphasis, e.g. *end up, cover up, fill out, pick up* or *pay off*.

(40) *I s= **filled out** all the[i:] (er) UCAS forms to decide where to go* (LOCNECB)

[27] The numbers of each group may vary slightly, as some PhV require different frames in a different context, based on their current semantic meaning.
[28] It should be noted, however, that *put on weight* is an echo repetition of the student after the teacher. This case is therefore disregarded in the quantitative analysis.
[29] In fact, that is the case with the majority of PhV.

PhV found in the LOCNECB corpus can be assigned to the semantic categories of activity, (*shut up*), communication (*bring up*), mentality (*find out*), occurrence (*run out*) and aspect (*keep on*).

Table 12 summarizes all the PPV that appear in the LOCNECB corpus:

Phrasal-prepositional verbs – LOCNECB	
MWV	**raw frequency**
Get on with	8
Go back to	7
Look forward to	5
Come out with	3
End up with	3
Fit in with	3
Stand up to	3
Come out of	2
Come up with	2
Make up for	2
Get away with	2
Keep up with	1
Look out for	1
Put up with	1
Stay out of	1
Get through to	1
Get down to	1
TOTAL	46

Table 12 – all phrasal-prepositional verbs in LOCNECB

Overall, there are 17 PPV types in LOCNECB which account for 46 tokens. The most frequent PPV is *get on with* ('have a good relationship'; 8 occurrences) followed by *go back to* ('return to an activity/state that used to happen/be'; 7 occurrences), and *look forward to* ('happily expect'; 5 occurrences)

(41) *...so although I **get on** really well **with** <first name of a girl>...* (LOCNECB)

(42) *Yeah a really good book so I read it first [] I must have been around fourteen [ ] and then I **went back to** it tried again* (LOCNECB)

Other PPV that are found in the corpora are, e.g. *put up with* (1 occurrence), *keep up with* (1 occurrence) or *stand up to* (3 occurrences). The most frequent particle forming PPV is *up*, whereas *with* is the most frequent preposition.

(43) *it's all really important I've got to put up with it* (LOCNECB)

Semantic categories of activity (e.g. *stand up to*), mentality (e.g. *put up with*) and causativeness (e.g. *end up with*) are the only ones represented in the data.

4.3. COMPARISON OF MWV IN LINDSEI_CZB AND LOCNECB

Both non-native, Czech speakers (of LINDSEI_CZB) and native speakers of English (of LOCNECB) are found to use all the researched MWV types, i.e. phrasal verbs, prepositional verbs and phrasal-prepositional verbs. A quantitative comparison of types in the corpora is illustrated in Table 13[30]. As the numbers show, both raw and normalized frequencies per 100000 tokens yield the same comparative results.

Type representation of MWV in the corpora				
Category	**LINDSEI_CZB**		**LOCNECB**	
	Raw frequency	Normalized frequency	**Raw frequency**	Normalized frequency
PhV	**77**	78	**155**	145
PrV	**80**	81	**73**	68
PPV	**13**	13	**17**	16
TOTAL	**170**	172	**245**	229

Table 13 – type representation in the corpora

The results show that the native speakers in LOCNECB use 245 (229)[31] types of the three MWV classes, over 170 (172) MWV than we find in the LINDSEI_CZB corpus. The difference (when both raw and normalized frequencies are considered) is significant at $p<0.05$. Thus, this should be taken as the first indicator of a greater MWV preference in LOCNECB over LINDSEI_CZB. In addition, the category of PhV is represented by more types in LOCNECB, i.e. 155 (145) PhV types over 77 (78) types in LINDSEI_CZB. The difference is significant at $p<0.0001$; there is therefore a big gap between the usage by non-native and native speakers pointing to the phenomenon of PhV underuse (and thus possibly avoidance) by non-native speakers. Native speakers also use more PPV types (i.e. 17 (16) over 13 (13) types); however, the difference is not statistically significant. Although non-native speakers use more of PrV

[30] The corpora were normalized to 100 000 words. Normalized frequency equitation used: number of a given MWV in a corpus ÷ number of words in its <B> corpora × 100 000, e.g. PrV in LINDSEI_CZB: 80 ÷ 98 814 × 100 000 = 80, 9601. The final normalized numbers were rounded to integers, i.e. 81 PrV in LINDSEI_CZB.
[31] Numbers in brackets denote normalized frequency of MWV.

types than native speakers, the difference is not statistically significant. It can be thus concluded that the usage of PrV (and also PPV) types is rather comparable than significantly different in terms of (non)nativeness; however, the use of PhV is clearly prevalent in the native corpus.

To paint the whole picture, token distribution of the corpora is presented in Table 14.

Token representation of MWV in the corpora				
Category	**LINDSEI_CZB**		**LOCNECB**	
	Raw frequency	Normalized frequency	**Raw frequency**	Normalized frequency
PhV	**191**	193	**681**	637
PrV	**398**	403	**426**	398
PPV	**33**	33	**46**	43
TOTAL	**622**	630	**1153**	1078

Table 14 – token representation in the corpora

The comparison of token results between PhV and PPV in LINDSEI_CZB and LOCNECB confines to the pattern which has been set by their types: there are more tokens of PhV (681 (637) vs 191 (193)) and PPV (46 (43) vs 33 (33)) in LOCNECB than in LINDSEI_CZB corpus. The difference between the quantity of PhV is significant (at $p<0.0001$) with both normalized and raw frequency, as there is over three times more tokens of PhV in the native corpus. When the raw frequency is considered, the token representation of PrV is in contrast to the type representation; whereas there are more types of PrV in LINDSEI_CZB, the types are represented by more tokens in LOCNECB. In other words, although there are more types of PrV in LINDSEI_CZB, there are more tokens of this class in LOCNECB. However, when normalized frequency is considered, it shows that the representation of PrV across the non-native corpus (403 tokens) is higher than in the native corpus (398 tokens). The difference between the PrV token frequency is, however, not statistically significant and it can be thus concluded that usage of PrV tokens is rather comparable across the corpora.

The raw number of PrV types in both the corpora is quite similar, however, this does not mean that they are the same PrV types in the corpora. In fact, the corpora share only 49 types of PrV, usually the most frequent ones, e.g. *think about, go through, look like* etc., the remaining 24 in LOCNECB (e.g. *look through, cope with,* and *rely on*) and 31 in LINDSEI_CZB (e.g. *laugh at, draw on,* and *reflect on*) differ based on the corpora. This may be the reason for the difference in tokens, as both groups used different PrV to a different extent.

The data partially confirm Hypothesis 1. Hypothesis 1 posits that the native speakers will use all MWV more frequently than the non-native speakers in this study. This is true when

the total numbers of MWV are considered[32]. As Table 13 and 14 show, this is also the case of raw token representation of all the MWV studied, i.e. PhV, PrV and PPV and of type representation of PhV and PrV. However, type representation of PrV does not conform to Hypothesis 1 as more PrV types are found in the non-native corpus. In addition, when the normalized frequency is considered, PrV tokens are actually more widespread in the non-native corpus. However, we see the usage of PrV and also PPV in terms of both types and tokens as rather comparable as the differences are not statistically significant. The main difference of MWV usage thus lies with the PhV type and token representation, which is a clear indicator of the different situation with the two groups of speakers.

Although the representation of PrV does not conform to Hypothesis 1; it does prove Hypothesis 2 as PrV class is the most frequent MWV (in respect to both type and token representation) in the LINDSEI_CZB corpus. The number of PrV tokens when compared to PhV and PPV tokens, is statistically significant at $p<0.0001$ again pointing to the PrV dominance in the LINDSEI_CZB corpus.

In order to answer Hypothesis 3, the corpora are compared for the exact number of tokens that appear in the data. Table 15 presents ten most frequent PhV in each of the corpora.

LINDSEI_CZB PhV		**LOCNECB PhV**	
MWV type	**raw frequency**	**MWV type**	**raw frequency**
Find out	18	Go back	50
Come back	17	Come back	42
Go on	11	Go out	27
Go back	9	Go on	26
Turn out	9	**Work out**	18
Go out	7	Get out	17
Pick up	7	Get up	16
Give up	4	End up	14
Go over	4	Get back	14
Sit down	4	Come up	13
TOTAL	90	**TOTAL**	237

Table 15 – top ten PhV in LINDSEI_CZB and LOCNECB

PhV that appear in Table 15 can be found in both the corpora, four of which appear at their top ten positions, i.e. *go back* (first and fourth position in LOCNECB and LINDSEI_CZB respectively), *come back* (second place in both), *go out* (third and sixth respectively) and *go on* (fourth and third respectively). These four PhV are the most frequent in LOCNECB and when compared to their LINDSEI_CZB synonyms, they are also represented by more tokens.

[32] This applies when raw and also normalized frequencies are considered

These ten PhV thus fully illustrate the difference between the state of PhV in the corpora, i.e. the ten most frequent types in LOCNECB amount to 237 tokens but to only 90 tokens in LINDSEI_CZB. The most frequent PhV in the student corpus is *find out* (18 occurrences), which has eight occurrences in the native corpus. *Find out* therefore seems to conform to the phenomenon of the so-called 'phrasal teddy-bears' in that it is more frequent (by 10 tokens) in the student than in the native corpus. In addition, it seems to be used two times in a sense that is not very native-like (sentences 44 and 45); in fact, the native speakers used it only in the meaning denoting *discover*

(44) *I think law is exciting you know you have to **find out** a set of rules to . (erm) . (em) . govern the society* (LINDSEI_CZB)

(45) *I did biology in English at my school so I was **found out** to be sufficient for this job* (LINDSEI_CZB)

The first case is close to denoting 'discover' in a sense that something has to be revealed and understood. However, the context suggests that the speakers rather meant 'learn' or 'establish'. In the second example, the PhV is actually overgeneralized in that it is used in context where a single-word *find* would suffice. It thus seems that the non-native speakers rely on this particular PhV more and use it in more senses than the native speakers in this study. The difference is also statistically significant at $p<0.05$. Moreover, single-word synonyms (i.e. *discover* and *realize*) were checked in order to find out whether the native speakers favour them over the PhV. This was proven not to be the case, as both *discover* and *realize* appear once in the whole corpus; on the contrary, there are five occurrences of *discover* and 34 occurrences of *realize* in LINDSEI_CZB. It thus seems that the native speakers have no need to use such an expression much in their speech, whereas the non-native speakers rely on it quite a lot, also when its one-word synonyms are concerned. *Find out* thus fulfils Hypothesis 3 in that student seem to overuse it as well as use it in non-native like contexts. The analysis also reveals that non-native speakers, besides overusing MWV, underuse certain relatively frequent MWV in LOCNECB. This is the case of *work out* which appears on the fifth position for native speakers (18 tokens and 12 speaker), but only four non-native speakers (i.e. four tokens) use it in their speech.

In order to discover, whether a certain MWV is not frequent only because a certain speaker favours the given MWV, Table 16 provides numbers of speakers (Sp) that use a given PhV. As there are 50 non-native speakers and 39 native speakers in this study, normalized

frequency of speakers (NF of Sp) per 50 speakers[33] is counted in order to compare uses of the same MWV in the corpora.

PhV LINDSEI_CZB				PhV LOCNECB			
MWV	**Sp**	**NF of Sp**	**Tokens**	**MWV**	**Sp**	**NF of Sp**	**Tokens**
Find out	11	11	18	Go back	18	23	50
Come back	12	12	17	Come back	22	28	42
Go on	7	7	11	Go out	14	18	27
Go back	6	6	9	Go on	17	22	26
Turn out	7	7	9	**Work out**	12	15	18
Go out	7	7	7	Get out	14	18	17
Pick up	4	4	7	Get up	10	13	16
Give up	4	4	4	End up	11	14	14
Go over	3	3	4	Get back	10	13	14
Sit down	3	3	4	Come up	12	15	13
TOTAL	64	64	90	**TOTAL**	140	179	237

Table 16 – speaker/token ratio of PhV in the corpora

Table 16 shows several interesting findings. *Come back* is the widest dispersed PhV in both of the corpora (although it is not the most frequent PhV in neither of them). The number of tokens in LOCNECB's top four positions drops quite dramatically if we consider only one token for one speaker for a given PhV, i.e. eighteen speakers (normalized to 23) use *go back* in LOCNECB, however, the overall number of tokens (due to several repetitions by speakers) has risen to the total of 50 occurrences. The PhV is also used at least two times per speaker (with the exception of two cases where the speakers used it only once); the maximum being six occurrences with one speaker. Further, the frequency was increased by one speaker in the case of *work out* (used by one speaker six times) in LOCNECB. Overall, it is clear that PhV are less dispersed in LINDSEI_CZB than in LOCNECB. When the normalized frequency of speakers is taken into account, we see that the top ten PhV are more widespread in LOCNECB than in LINDSEI_CZB. There is also a clear decrease in the number of speakers, starting with *pick up* down-wards in LINDSEI_CZB. Therefore, the normalized frequency of speakers shows a wider dispersion of PhV in LOCNECB which coincides with there being three times more PhV tokens in the native corpus. In summary, PhV are thus clearly underused by non-native speakers.

[33] e.g. normalized frequency for LOCNECB is counted as: a number of speakers per type ÷39 x 50.

Table 17 presents ten most frequent PrV in the respective corpora.

LINDSEI_CZB PrV		**LOCNECB PrV**	
MWV type	**raw frequency**	**MWV type**	**raw frequency**
Talk about	40	Look at	55
Think about	38	Think about	29
Look at	36	Go into	27
Look like	21	Look like	22
Go for	20	Talk about	20
Listen to	14	Think of	17
Talk to	12	Come from	16
Focus on	11	Apply for	13
Say to	10	**Say to**	12
Wait for	9	**Write to**	12
TOTAL	211	**TOTAL**	223

Table 17 – top ten PrV in LINDSEI_CZB and LOCNECB

PrV that appear in Table 17 can be found in both corpora, five of which appear at their top ten positions, i.e. *look at* (first and third position in LOCNECB and LINDSEI_CZB respectively), *think about* (second place in both), *look like* (fourth place both), *talk about* (fifth and first position respectively) and *say to* (ninth position in both). The token representation of PrV is quite comparable across the corpora, in that the difference in frequency on top positions is not as striking as it is with PhV, i.e. there is only 12 tokens difference between PrV as opposed to 147 token difference with PhV). Also, the number of PrV that appear at least two times in the respective corpora is identical: 48 in LOCNECB a 48 in LINDSEI_CZB. If the PrV types are studied more closely, they appear to oscillate on a scale between being more transparent to being rather more idiomatic. To exemplify, e.g. *talk about, think about, listen to, apply for, spy on*, or *thank for* that are found in either or both corpora, seem to incline to the more transparent side of the scale, the constituents are both quite transparent, however, the particle is governed by the verb and thus they create a functional whole. On the other hand, e.g. *look after (~care), go after (~chase), come across (~find unexpectedly), turn into (~change)* or *hit on (~seduce*) are rather positioned on the more idiomatic side; it is quite impossible to deduce the meaning from the two constituents only.

As has been already mentioned, the number of the PrV *talk about* is reduced to 40. It is the most frequently used PrV by non-native speakers who often start their interviews in that way (even when not repeating the PrV after the interviewer). Although not used in different

contexts, thanks to its double number of occurrences (i.e. 40 in LINDSEI_CZB vs 20 in LOCNECB) it can also be seen as a comforting and overused phrasal teddy-bear. Its dispersion among speakers in the corpora, especially when we consider normalized frequency is however, rather comparable. Non-native speakers also heavily underuse certain PrV, e.g *go into* (27 occurrences and 13 speakers in LOCNECB against one occurrence in LINDSEI_CZB). *Think about* is the second most frequent PrV in both corpora, appearing more frequently in LINDSEI_CZB (i.e. 38 vs 29 tokens). Other PrV with *think*, *think of*, is far less popular in the non-native corpus (7 vs 17 tokens). It was hypothesized that this discrepancy may be caused by preference for *think about* over *think of* in LINDSEI_CZB. However, it has not proven to be the case, as we find only one example where *think of* would be more suitable than *think about* (sentence 46), for the reason of denoting 'remember'. In other cases, the use of *think about* is seen as founded.

(46) *I can't* ***think about*** *anything better now* (LINDSEI_CZB)

Listen to does not appear in the native corpus, although it is quite frequent in LINDSEI_CZB, i.e. 14 tokens vs none. Also, *turn to* (3 occurrences) and *turn into* (8 occurrences) when combined, appear 11 times in the non-native corpus and only once in the corpus of native speakers. It thus seems that *listen to* as well as *turn to/turn into* (both denoting 'change') are additional examples of phrasal teddy-bear phenomenon. Also, whereas the one example of *turn to/into* in LOCNECB expresses a change of an actual physical being:

(47) *He got t=* ***turned into*** *a vampire which he didn't like very much* (LOCNECB)

the examples in LINDSEI_CZB also denote a sort of figurative sense, which is however, natural also in the native context

(48) *The movie then* ***turns into*** *a[ei] (eh) basically a very solitary (eh) experience* (LINDSEI_CZB)

The non-native speakers use *turn to/into* in various senses that are seen as native-like, however, only one native speaker actually uses it in our data. It thus seems that non-native speakers do not diverge from the natural uses, however, they do seem to give it more space in their speech. When the speaker/token ratio is considered we see that four speakers use *turn into* and two speakers employed *turn to*, combining to six speakers in LINDSEI_CZB vs one (normalized to one) in LOCNECB. After closer inspection, such a quite high number of tokens of *turn into* is caused especially by one speaker, who used it four times during her/his interview. Thus, phrasal teddy-bears are not only a characteristic of the non-native group but also of non-native (and also native) individuals. This seems to be the case with many recurring PrV in one speaker file.

The PrV, or any MWV that recur in a speech of one student often follow one another in close proximity; when a given MWV is used once, is at hand and is thus probably easier to use again.

Turn into is not the only PrV where the speaker/token ratio differs and influences the frequency of PrV. Table 18 provides differences between the number of PrV tokens and of speakers that use them. It also gives normalized frequency of speakers.

LINDSEI_CZB PrV				**LOCNECB PrV**			
MWV	**Sp**	**NF of Sp**	**Tokens**	**MWV**	**Sp**	**NF of Sp**	**Tokens**
Talk about	20	20	40	Look at	29	37	55
Think about	18	18	38	Think about	19	24	29
Look at	22	22	36	Go into	13	17	27
Look like	14	14	21	Look like	15	19	22
Go for	9	9	20	Talk about	14	18	20
Listen to	10	10	14	Think of	16	21	17
Talk to	11	11	12	Come from	12	15	16
Focus on	7	7	11	Apply for	8	10	13
Say to	8	8	10	**Say to**	9	12	12
Wait for	7	7	9	**Write to**	5	6	12
TOTAL	126	126	211	**TOTAL**	140	179	223

Table 18 – speaker/token ratio of PrV in the corpora

Table 18 shows that neither PrV on the top position has a speaker/token correspondence. In fact, with the case of *talk about* in LINDSEI_CZB, twenty speakers use it to various degrees, i.e. eight speakers use it more than once. In fact, the eight speakers use up 28 tokens out of 40. In one case, one speaker uses the PrV eight times. Clearly, the dispersion across the texts is not as wide as the raw token frequency may have suggested, although it is still quite high, compared to other PrV (as 20 out of 50 speakers actually use it in their speech). *Look at* is the most widespread PrV in LOCNECB (extendedly the most widespread MWV overall), although one speaker used it eight times. When the normalized frequency is considered, the number of speakers also indicates its overall popularity. *Look at* is also the most dispersed PrV (and MWV) in LINDSEI_CZB: 22 speakers share 36 tokens; out of these, eight speakers share 22 tokens.[34] These results confirm the findings of Biber et al (1999: 416) who claim that *look at* is the most frequent PrV (mainly in conversation and fiction). Unlike in the case of PhV, where there is a

[34] To be more specific, one speaker used it six times, two speakers used it three time, and five speakers used it two times.

clear difference in the width of dispersion in the two corpora, the dispersion on the top positions with PrV is definitely more equal.

Among other PrV that are not in top ten, protrudes the PrV *be into*. *Be into* appears seven times in LINDSEI_CZB and two times in LOCNECB. The frequency of *be into* in LINDSEI_CZB is however, enhanced by one speaker that repeated it four times (almost in a row); frequency is also enhanced in LOCNECB with *go into*: one speaker used it seven times during his/her interview. Although there are more cases, when the speaker/token ratio does not equal, there are some cases (although more scarce) where one token corresponds to one speaker, e.g. *get into* (8 speakers/occurrences), *go through* (5) , *know about* (4), and *base on* (3) in LINDSEI_CZB.

PrV types on the top nine positions appeared at least ten or more times in LINDSEI_CZB, compared to 14 types in LOCNECB. On the other hand, there is a great discrepancy between the highest numbers of tokens when PhV are concerned. There are only three PhV that appear more than ten times in LINDSEI_CZB, whereas there are fourteen such PhV in LOCNECB. This again seems to point to PhV, if not avoidance, then underuse by the non-native students

Lastly, Table 19 compares the use of ten most frequent PPV in the corpora.

PPV LINDSEI_CZB				**PPV LOCNECB**			
MWV	**Sp**	**NF of Sp**	**Tokens**	**MWV**	**Sp**	**NF of Sp**	**Tokens**
Look forward to	7	7	12	Get on with	5	6	8
Go back to	2	2	5	Go back to	6	8	7
Come up with	2	2	3	Look forward to	5	6	5
Get away from	2	2	2	Come out with	2	3	3
Get back to	1	1	2	End up with	2	3	3
Get on with	2	2	2	Fit in with	2	3	3
Catch up with	1	1	1	Stand up to	1	1	3
Come back to	1	1	1	Come out of	1	1	2
End up with	1	1	1	Come up with	2	3	2
Get around to	1	1	1	Make up for	2	3	2
TOTAL	20	20	30	**TOTAL**	28	37	38

Table 19 – top ten PPV in the corpora; with speaker token ratio

Look forward to is the most frequent PPV in LINDSEI_CZB, with 12 occurrences. It is also PPV with widest dispersion in this corpus, as seven non-native speakers use it in their speech. The difference in native and non-native use (12 vs 5 tokens) points to it being a phrasal teddy-bear of non-native speakers. Conversely, *get on with* (8 occurrences and 5

(normalized 6) speakers in LOCNECB and two in LINDSEI_CZB) is rather underused by the non-native speakers in their utterances.

Together with *go back to*, *look forward to* is quite in contrast with the remaining PPV in the non-native corpus: the rest of PPV are very sporadic. Actually, half of the PPV only occur once in the whole non-native corpus. Although there are more variable and also overall more frequent PPV in LOCNECB, they cannot match the frequency of the remaining MWV categories. Whereas there are three PPV in LOCNECB: *get on with, go back to* and *look forward to* which are relatively (in regard to other PPV only) widespread, there is only one such PPV in LINDSEI_CZB: *look forward to*. In addition, *get on with* is with eight tokens the most frequent PPV in LOCNECB, however, this frequency is boosted by one speaker that used it four times in his speech. Because of this, *go back to* should be seen as the PPV with the widest dispersion, as six speakers (normalized to eight when compared with LINDSEI_CZB) used seven tokens (meaning one speaker repeated it once), followed by *look forward to* having five speakers = five tokens correspondence in LOCNECB. Although some PPV are represented by two or more tokens, the frequency is influenced by only one speaker, i.e. *get back to* (LINDSEI_CZB), *stand up to*, and *come out of* (LOCNECB). The position of *stand up to* is especially influenced by this, as the speaker uses it three times.

5. CONCLUSIONS

MWV proved to be a complicated and elusive field burdened by several issues. The academic articles show that learning this specific type of verb is a difficult process because there is not a clear consensus on how to approach MWV neither as a teacher nor a student. However, the articles agree that it is necessary to focus first on the most frequent MWV and teach them in as much natural setting as possible (Trebits, 2009: 477). This way, the students can recognize them outside of the classroom and acknowledge their use and correctness and later on also expand their MWV vocabulary bank. It is also beneficial to introduce the structures early on in the learning stage in order to make them a recognizable part of language learning and thus avoid their stigmatization (Yang & Hseih, 2010: 6). There is a need for introducing not only the form, but also the possible different semantic senses that are based on context (Trebits, 2009: 477). As the Analysis shows, this is the case with many PhV, e.g. *work out, go on*, or *put on.* Many authors focus their attention on the avoidance of this verb category caused by e.g. arbitrary exposure, low proficiency or indirect interference (Dagut & Laufer, 1985: 78, Liao & Fukuya, 2002: 90), as MWV are often neglected in favour of single-word synonyms that are probably seen as more accessible and less complicated in both form and sense (Yasuda, 2010: 251). It is therefore important to incorporate this structure systematically, from the early stages of learning and in many contexts.

The avoidance of MWV is indirectly confirmed by the results in the Analysis. Table 20 gives the type and token frequencies of all the classes in both corpora. MWV on a grey background are the most frequent MWV class in the respective corpus. Green colour signals the most numerous type representation across the corpora. The most numerous token representations are marked by orange colour.

MWV types and tokens in the corpora				
	LINDSEI_CZB		**LOCNECB**	
MWV class	**Types**	**Tokens**	**Types**	**Tokens**
PhV	**77** (78)	**191** (193)	**155** (145)	**681** (637)
PrV	**80** (81)	**398** (403)	**73** (68)	**426** (398)
PPV	**13** (13)	**33** (33)	**17** (16)	**46** (43)
TOTAL	**170** (172)	**622** (630)	**245** (229)	**1153** (1078)

Table 20 – MWV types and tokens in the corpora

The biggest difference in use is seen with PhV. We found 77 (78) types and 191 (193) tokens in the non-native corpus as opposed to 155 (145) types and 681 (637) tokens in the native corpus. The difference in both types and tokens is statistically significant at $p<0.0001$ when

both raw and normalized frequencies are considered. Clearly, non-native speakers underuse PhV to an extent which is unnatural in the native context. There are also more PPV types and tokens in the non-native corpus, although the difference is not statistically significant; however, it should be noted that the number of PPV tokens is mainly enhanced by PPV *look forward to*, which appears 12 times in LINDSEI_CZB. We found more PrV types in LINDSEI_CZB than in LOCNECB when the raw as well as normalized frequencies are considered. There are more PrV tokens in LOCNECB when raw frequency is considered; however, when normalized frequency is taken into account, PrV are more dispersed in the non-native corpus. The difference is not statistically significant neither when PrV type/token difference nor frequency types are considered. The usage of PrV is thus rather comparable between the two groups of speakers. Hypothesis 1 (presupposing that native speakers will use all the MWV more than the non-native speakers) is thus only partly confirmed. It is confirmed when the overall numbers are considered, non-native speakers are found to use 170 (172) types and 622 (630) tokens of MWV, whereas native speakers used 245 (229) types of MWV and 1153 (1078) tokens. The difference between both types and tokens is statistically significant at $p<0.0001$. In addition, the natives use more tokens of all the main classes and they also employ more PhV and PPV types. However, the non-native speakers use more PrV types and when the normalized frequency is factored in the data, PrV tokens are more widespread in the non-native data. However, the only statistically significant difference is in the use of PhV which are clearly underused by non-native speakers, the use of the remaining two MWV categories is rather comparable, especially in the case of PrV.

These findings allow us to confirm Hypothesis 2 (predicting that PrV will be the most used class of MWV by non-native speakers). 80 types of PrV were found in LINDSEI_CZB as opposed to 77 PhV and 13 PPV. Although the difference between PhV and PrV types is not statistically significant, the difference between their tokens is significant at $p<0.0001$; there are 191 PhV tokens and 398 PrV tokens.

To prove or disprove Hypothesis 3 (regarding favoured MWV, so called "phrasal teddy-bears"), the data are analysed in more depth. Table 21 summarizes five most frequent MWV found in each of the categories and in both corpora. The numbers in brackets denote normalized frequency of speakers (to 50). There are 50 non-native speakers and 39 native speakers in this study.

MWV in LINDSEI_CZB and LOCNECB					
MWV	**Speakers**	**Tokens**	**MWV**	**Speakers**	**Tokens**
PrV in LINDSEI_CZ			**PrV in LOCNECB**		
Talk about	20 (20)	40	Look at	29 (37)	55
Think about	18 (28)	38	Think about	19 (24)	29
Look at	22 (22)	36	Go into	13 (17)	27
Look like	14 (14)	21	Look like	15 (19)	22
Go for	9 (9)	20	Talk about	14 (18)	20
TOTAL	*83*	*155*	*TOTAL*	*90*	*153*
PhV in LINDSEI_CZB			**PhV in LOCNECB**		
Find out	11 (11)	18	Go back	18 (23)	50
Come back	12 (12)	17	Come back	22 (28)	42
Go on	7 (7)	11	Go out	14 (18)	27
Go back	6 (6)	9	Go on	17 (22)	26
Turn out	7 (7)	9	**Work out**	12 (15)	18
TOTAL	*43*	*64*	*TOTAL*	*83*	*163*
PPV in LINDSEI_CZB			**PPV in LOCNECB**		
Look forward to	7 (7)	12	Get on with	5 (6)	8
Go back to	2 (2)	5	Go back to	6 (8)	7
Come up with	2 (2)	3	Look forward to	5 (6)	5
Get away from	2 (2)	2	Come out with	2 (3)	3
Get back to	1 (1)	2	End up with	2 (3)	3
TOTAL	*14*	*24*	*TOTAL*	*20*	*26*

Table 21 – top five MWV in each of the classes in both corpora

Talk about, find out and *look forward to* are the most promising examples of the phrasal teddy-bear phenomenon in Table 21, accompanied by also *listen to* (14 occurrences in LINDSEI_CZB vs none in LOCNECB) and a merge of *turn into* and *turn to* (11 vs one occurrence respectively). With these MWV (with the exception of *find out*) it is only the increased frequency that points to the direction of phrasal teddy-bears. With *find out*, however, it is the increased frequency as well as overgeneralization in contexts that are not completely native-like that completely fulfil this concept. Therefore, Hypothesis 3 is confirmed by at least one, extendedly by five MWV, i.e. it was proven that the non-native students seem to favour *find out* more than is native-like and they also overgeneralize them and thus use them in non-native like contexts. Individual speakers (both native and non-native) also seem to have their teddy-bears, e.g. *be into* (four occurrences and one speaker) in LINDSEI_CZB or *go into*

(seven occurrences) in LOCNECB. Conversely, it is also the case that non-native speakers tend to underuse certain MWV that are quite frequent with native speakers, e.g. *work out,* which is the fifth most frequent PhV in LOCNECB (18 tokens and 12 (15) speakers), is used only four times (and by four speakers) in LINDSEI_CZB, *go into* (27 occurrences and 13 (17) speakers in LOCNECB and only one on LINDSEI_CZB), or *get on with* (8 occurrences and 5 (6) speakers as opposed to two in LINDSEI_CZB).

Look at is the most frequent PrV (as well as MWV) in LOCNECB and third most frequent in LINDSEI_CZB. In addition, it is the widest dispersed MWV across the corpora; it was used by 29 (normalized to 37) and 22 (22) speakers, respectively. Thus, not only is its popularity supported by number of tokens, but also by the number of speakers. The dispersion across the corpora also supports the idea of speakers' repetitions: if a MWV type is used at all, it is more often than not used more than once. In consequence, it also helps to paint a bigger picture that although a certain MWV is numerous in frequency, it may not be as widely dispersed, e.g. *go back, come back* in LOCNECB or *think about,* or *go for* in LINDSEI_CZB. The frequency of *look at* in LOCNECB and wide dispersion in both corpora also coincide with findings of Biber et al, (1999: 416) who claim that *look at* is the most frequent PrV (mainly in conversation and fiction).

Data in Table 21 present another feature of MWV that can be seen across the whole corpora. MWV are said to be marked by idiomaticity of their meaning (which is one of the deemed reasons for their avoidance). It is clear only by the handful of examples in Table 21 that idiomaticity cannot be seen as one point in space, but rather a scale from almost transparent to highly idiomatic. This scale is applicable to PhV but mainly in PrV identification. To exemplify with PrV, e.g. *come back, listen to, spy on*, or *thank for* are rather transparent in that the constituents both hold on to their original meanings, however, the particle is governed by the verb and thus they create a functional whole. On the other hand, e.g. *go for (~choose), look after (~care),* or *hit on (~seduce*) are rather positioned on the more idiomatic side; it is quite impossible to deduce the meaning from the two constituents only. Thus, we should not rely excessively on the idiomatic aspect (although it is an important part of mainly PhV identification) and focus also on what the V + particle offers in terms of collocability. It is less prominent with PhV, but we can see the need for a scale in cases such as *come back* (inhabiting the more transparent side) on one hand and *give up (~quit)* or *take up* (*~begin*) (belonging to the idiomatic part) on the other.

The analysis also posed several difficulties. Due to extensive manual analysis, we allow for human error in MWV identification. Another difficulty arises with the identification of

MWV, especially with differentiation between MWV and FC. This has proven especially difficult with PrV as academic sources and dictionaries largely differ in their interpretation. It would be beneficial for further research to focus on the actual situation of teaching and learning (i.e. identify means through which MWV are learnt, if and consequently how are MWV incorporated in the classroom), benefiting from the current data presented here about underuse and overuse of MWV. In addition, further research focusing on variables other than mother tongue of the non-native speakers (e.g. proficiency level, other languages known, length of a stay at English speaking country, or onset of student's English learning) could bring a more complex insight into the field in the context of Czech speakers, possibly shading a light at the most important factors for English MWV acquisition.

6. REFERENCES

Abdel Salam El-Dakhs, D (2016) "The Lexical Knowledge and Avoidance of Phrasal Verbs: The Case of Egyptian Learners of English". *International Journal of Applied Linguistics and English Literature*, vol. 5., no.1, 132-144. Available at: <http://journals.aiac.org.au/index.php/IJALEL/article/view/1970>. [Last accessed 8 February 2020]

Barekat, B. & Baniasady, B. (2014). "The Impact of Phrasal Verb Avoidance on the Writing Ability of the University EFL Learners". *Procedia – Social and Behavioral Sciences* 98. Elsevier Ltd. 343-352. Available at: https://www.sciencedirect.com/science/ [Last accessed 9 February 2020]

Biber, D. et al. (1999) *Longman Grammar of Spoken and Written English*. London: Longman

Cambridge Dictionary (2020) Cambridge University Press. Available at: https://dictionary.cambridge.org/ [Last accessed 27 April 2020]

Collins Dictionary (2020) Harper Collins Publishers. Available at: https://www.collinsdictionary.com/dictionary/english [Last accessed 2 May 2020]

Dagut, M., Laufer, B. (1985) "Avoidance of Phrasal Verbs – A Case of Constrastive Analysis". *Studies in Second Language Acquisition*, vol. 7, No.1 73-80. Cambridge University Press.

De Cock, S. (2006). "Learners and Phrasal Verbs". *MED Magazine*, issue 35. Available at: http://www.macmillandictionaries.com/MED- [Last accessed 8 February 2020]

De Cock, S., Gilquin G., Granger S. (eds.) (2010) *Louvain International Database of Spoken English Interlanguage (LINDSEI)*. Belgium: Presses universitaires de Louvain.

Dušková, Libuše et al. (2009) ***Mluvnice současné angličtiny na pozadí češtiny***. Electronic edition. Ústav anglického jazyka a didaktiky, FF UK. Available at http://www.mluvniceanglictiny.cz [Last accessed 13 December 2019].

Ellis, N. C. (2012) "Formulaic Language and Second Language Acquisition: Zipf and the Phrasal Teddy Bear." *Annual Review of Applied Linguistics*, 32, 17-44. Cambridge University Press.

Gardner D., Davies M. (2007). “Pointing out Frequent Phrasal Verbs: A Corpus-Based Analysis” *TESOL Quarterly*, vol. 41, No. 2. Available at: https://www.jstor.org/stable/40264356? [Last accessed 27 January 2020]

Gilquin, G. (2015) “The Use of Phrasal Verbs by French-speaking EFL Learners. A Constructional and Collostructional Corpus-based Approach”. *Corpus Linguistics and Linguistic Theory* 11(1). Available at: https://www.researchgate.net/publication/272560957_The_use_of_phrasal_verbs_by_French [Last accessed 31 January 2020]

Ghabanchi Z. & Goudarzi E. (2012). “Avoidance of Phrasal Verbs in Learner English: A Study of Iranian Students”. *World Journal of English Language*. Vol. 2, no. 2. Available at: https://pdfs.semanticscholar.org/326b/ [Last accessed 8 February 2020]

Hasselgård, H. (2019). "Phraseological teddy bears". *Corpus Linguistics, Context and Culture*, 15, 339-362.

Houshyar S. & Talebinezhad, M (2013). “Study on Avoidance Behaviour among Persian EFL learners: Phrasal Verbs in Focus”. *Greener Journal of Educational Research*. Vol. 3 (6). Available at: https://pdfs.semanticscholar.org/3873/ [Last accessed 8 February 2020]

Liao Y. D. & Fukuya Y.J. (2002) “Avoidance of Phrasal Verbs: The Case of Chinese Learners of English”. *Second Language Studies*. 20(2) Available at: https://www.semanticscholar.org/paper/Avoidance- [Last accessed 8 February 2020]

Nassaji, H., Tian J. (2010) “Collaborative and Individual Output Tasks and their Effects on Learning English Phrasal Verb” *Language Teaching Research*, vol. 14, issue 4. Available at: https://doi.org/10.1177/1362168810375364 [Last accessed January 29, 2020]

Quirk, R. et al. (1985) *A Comprehensive Grammar of the English Language*. London: Longman

Rundell, M et al. (ed.) (2020) Macmillan Dictionary. Macmillan Education Limited. Available at: https://www.macmillandictionary.com/ [Last accessed April 27 2020]

Summers, D et al. (dir.) (2000) *Longman Phrasal Verbs Dictionary*. England: Pearson Education Limited.

Trebits, A (2009) *The most frequent phrasal verbs in English language EU documents – A corpus-based analysis and its implications. System 37.* Elsevier Science. Available at:

https://www.sciencedirect.com/science/article/pii/S0346251X0900061X [Last accessed January 18, 2020]

Vendler, Z. (1957). "Verbs and Times". *The Philosophical Review*. Vol. 66, no. 2. 143-160. Available at: https://www.jstor.org/stable/2182371 [Last accessed 9 February 2020]

Wierszycka, J (2013) "Phrasal Verbs in Learner English: A Semantic Approach. A Study Based on a POS tagged spoken corpus of learner English" *Research in Corpus Linguistics.* Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.667.4501&rep=rep1&type=pdf [Last accessed: 30 January 2020]

Yasuda, S (2010) "Learning Phrasal Verbs Through Conceptual Metaphors: A Case of Japanese EFL Learners". *TESOL Quarterly*, vol. 44, No. 2. Available at: https://www.researchgate.net/profile/Sachiko_Yasuda2 [Last accessed: 31 January, 2020]

Yang, A & Hseih S. (2010). "Conceptual metaphor awareness on English phrasal verbs teaching and learning for adolescents in Taiwan". Available at: https://www.semanticscholar.org/paper/Conceptual-metaphor-awareness [Last accessed 1 February 2020]

You, Y. S. (1999) "Avoidance Phenomena of Phrasal Verbs by Korean Learners of English". *English Teaching*. 54(3). 135-155

Sources

Gráf, T. (2017). *LINDSEI_CZ: korpus spontánní mluvené angličtiny pokročilých mluvčích.* Ústav Českého národního korpusu FF UK, Praha 2017.

Louvain Corpus of Native English Conversation (2004) (LOCNEC).

SketchEngine Available from the SketchEngine web application: https://app.sketchengine.eu/#dashboard?corpname=user%2Fdivisovkl%2Flocnec_onlyb [Last accessed May 20 2020]

7. RESUMÉ

Diplomová práce se zabývá tématem používání víceslovných sloves (VSS) v řeči nerodilých (českých) a rodilých mluvčích angličtiny. Zaměřuje se hlavně na rozdíly v použití tří hlavních typů VSS, tj. slovesa frázová, slovesa předložková a slovesa předložková frázová. Práce je založená na předpokladu, že nerodilí mluvčí angličtiny (tj. univerzitní studenti na úrovní B2–C2) budou používat méně víceslovných sloves než rodilí mluvčí. Předpoklad je podpořen akademickými články zabývající se touto problematikou. Články se shodují na tom, že studenti ani učitelé nemají dostupný jednotný přístup ke studiu VSS, která jsou tak prezentována nahodile a v abstraktní podobě, načež právě prezentace VSS v konkrétních kontextech by ve velké míře přispěla k akvizici těchto sloves (Trebits, 2009: 477). Zároveň také dokazují, že pro studenty jsou VSS především obtížná kvůli jejich neexistenci v mateřském jazyce studenta (Gilquin, 2015: 10) či netransparentnosti jejich formy a významu (Yasuda, 2010: 251). V důsledku těchto nejasností, které se s VSS pojí, mají studenti tendenci se víceslovným slovesům vyhýbat (viz Dagut & Laufer, 1985; Liao & Fukuya, 2002 nebo Barekat & Baniasady, 2014).

Teoretická část nejdříve obecně prezentuje kategorii VSS, tj. jejich zasazení v kontextu slovesné kategorie, jejich formu i sémantiku. Dále jsou dopodrobna představeny jednotlivé typy VSS (frázová, předložková, předložková frázová). Teoretické zdroje se ve většině případů shodují na jejich rozdělení podle částice, která se s příslušným slovesem pojí. Frázové sloveso je tvořeno spojením slovesa a adverbiální částice, předložkové sloveso slovesem a předložkou a předložkové frázové sloveso adverbiální částicí a předložkou (Biber et al., 1999: 403). Při spojení slovesa s předložkovou částicí hraje také roli výsledná idiomatičnost vytvořeného frázového slovesa (ibid.), zatímco u předložkových sloves je nejdůležitějším faktorem syntaktická provázanost (Quirk et al., 1985: 1156). Některá víceslovná slovesa mohou také fungovat jako sloveso + částice ve volné kombinaci. Kombinace těchto dvou prvků na sebe není sémanticky nebo syntakticky vázaná (ibid.). V některých případech je také nutné od sebe odlišit víceslovná slovesa, která mohou fungovat jako frázová nebo jako předložková slovesa, např. *call on* (předložkové sloveso) vs. *call up* (frázové sloveso). V těchto případech se k rozlišení používají jak sémantické, tak hlavně syntaktické testy (např. postavení předmětu, možnost umístění adjunktu mezi sloveso a částici atd.) (ibid.: 1167).

V druhé části teorie jsou pak představeny tři hlavní témata z oblasti VSS, která se nejčastěji objevují v akademických článcích, tj. nejčastější VSS v různých druzích textů a v komunikaci (např. Gardner & Davies, 2007), přístupy k učení (se) VSS (např. Yang & Hseih, 2010) a výzkum a důvody pro vyhýbání se těmto slovesům (např. De Cock, 2006).

Metodologická část nejprve představuje tři hlavní hypotézy této práce, tj. nerodilí mluvčí budou používat méně VSS než rodilí mluvčí, nerodilí mluvčí budou více preferovat předložková slovesa (díky jejich větší transparentnosti) než zbylé typy, a budou mít zároveň v oblibě VSS, která jsou méně častá v rodilém korpusu a která (případně) budou používat i v kontextech, která nejsou typická pro rodilé mluvčí. Hypotézy jsou ověřované na dvou korpusech. Jedním z nich je žákovský korpus LINDSEI_CZ, který obsahuje rozhovory s padesáti českými studenty. Druhým korpusem je LOCNEC, který je referenčním korpusem LINDSEI, sestávající z rozhovorů se studenty s angličtinou jako prvním jazykem[35]. Korpus LINDSEI_CZ je jedním z komponentů mateřského korpusu LINDSEI, který dnes sestává z více jak jedenácti světových jazyků. LINDSEI_CZ a LOCNEC mají identickou strukturu (tj. rozhovory sestávají z úvodního monologu na dané téma, který přechází v diskuzi nad studentovými koníčky či plánovanou budoucností a jsou zakončeny převyprávěním příběhu na základě série obrázků).

Další část metodologie se již zaměřuje na zpracování samostatných dat. Značkování dat bylo provedeno v programu Sketch Engine. Data z LINDSEI_CZ a LOCNEC jsou rozdělena na tři sub-korpusy: celé korpusy, korpusy s odpověďmi pouze studentů (LINDSEI_CZB a LOCNECB) a korpusy s odpověďmi pouze tazatelů. Důvodem pro rozdělení celých korpusů je jednodušší práce s hlavními daty, tj. výpověďmi studentů. K vyhledání VSS jsou použity dva hlavní dotazy:

[tag="V.*"][tag="N.*"|tag="PP.*"|tag="DT"|tag="J.*"]{0,3}[tag="RP"|tag="RB.?"|tag="IN"]

[tag="V.*"][tag="N.*"|tag="PP.?"|tag="DT"|tag="J.*"]{0,3}[tag="RP"|tag="RB.?"][tag="IN"][36]

První dotaz je použit k vyhledání frázových a předložkových sloves, druhým dotazem jsou vyhledávána slovesa předložková frázová. Dotazy našly celkem 8559 možných tokenů VSS v LINDSEI_CZB a 10354 možných VSS tokenů v LOCNECB. Bylo již od počátku zjevné, že ne všechny nalezené kombinace mohou fungovat jako VSS. Po manuální selekci, konzultaci s gramatikami a slovníky se počet VSS snížil na finálních 622 tokenů (tj. 170 typů) v LINDSEI_CZB a 1153 tokenů (tj. 245 typů) v LOCNECB. Tato VSS byla roztříděna do třech hlavních typů VSS, zejména podle částice, idiomatičnosti a za pomoci identifikačních textů, gramatik a slovníků. Dále jsou v metodologii popsané některé limity této práce, např. velký objem dat či problémy při klasifikaci.

[35] Rozhovorů z LOCNEC je celkem třicet devět.

[36] Dotazy hledají sloveso, které je následované buďto částicí, adverbiem nebo předložkou (první dotaz) nebo částicí nebo adverbiem a předložkou (druhý dotaz). Mezi slovesem a částicí se také může, ale nemusí objevit substantivum, zájmeno, člen nebo adjektivum.

Analytická část nejdříve představuje výsledky z pohledu jednotlivých korpusů, v další části pak porovnává VSS v obou korpusech. Tabulka 22 shrnuje výskyt typů a tokenů VSS v jednotlivých kategoriích. Čísla napravo v závorkách představují počet typů a tokenů při normalizované frekvenci, zatímco čísla nalevo vyjadřují reálnou frekvenci typů a tokenů v korpusech.

	LINDSEI_CZB		**LOCNECB**	
Kategorie VSS	**Typy**	**Tokeny**	**Typy**	**Tokeny**
Frázová VSS	**77** (78)	**191** (193)	**155** (145)	**681** (637)
Předložková VSS	**80** (81)	**398** (403)	**73** (68)	**426** (398)
Předložková frázová VSS	**13** (13)	**33** (33)	**17** (16)	**46** (43)
SOUČET	**170** (172)	**622** (630)	**245** (229)	**1153** (1078)

Tabulka 22 – počet typů a tokenů v jednotlivých kategoriích VSS

Šedá barva v tabulce znázorňuje nejčastější VSS v daném korpusu. Zatímco v LOCNECB jsou nejčastější typem frázová slovesa (tj. 155 typů a 681 tokenů), v LINDSEI_CZB jsou nejčastější slovesa předložková (tj. 80 typů a 398 tokenů). Zelená a oranžová barva znázorňuje nejvyšší frekvenci dané kategorie (tj. typ nebo token) v porovnání obou korpusů. Je zjevné, že v LOCNECB je více typů sloves frázových a předložkových frázových, zatímco předložkových typů je nejvíce v LINDSEI_CZB. Slovesa frázová a předložková frázová jsou také reprezentována více tokeny v LOCNECB než v LINDSEI_CZB, zatímco při přihlédnutí k normalizované frekvenci je více předložkových tokenů v LINDSEI_CZB. Předpoklad, který byl formulován v Hypotéze 1 (tj. nerodilí mluvčí budou používat méně VSS než rodilí mluvčí) se tedy potvrdil jen zčásti. Potvrdilo se, že nerodilí mluvčí používají obecně méně VSS (jak typů tak tokenů) než rodilí mluvčí (rozdíl je statisticky významný, $p<0{,}0001$). Rodilí mluvčí také používají výrazně více frázových sloves (rozdíl jak u tokenů, tak typů je statisticky významný při $p<0.0001$) a používají také více typů a tokenů sloves předložkových frázových (rozdíl však není statisticky významný). Hypotézu nepotvrzuje použití předložkových sloves v LINDSEI_CZB.: při zvážení normalizované frekvence nacházíme v LINDSEI_CZB více typů i tokenů těchto sloves v porovnání s rodilým LOCNECB. Rozdíl v použití typů ani tokenů však není statisticky významný při $p<0.05$. Zdá se tedy, že nejmarkantnějším ukazatelem rozdílu použití jsou frázová slovesa, která jsou zjevně výrazně méně používaná nerodilými mluvčímu. Rozdíl v použití zbylých dvou kategorií VSS je naopak spíše srovnatelný, zejména v případě předložkových sloves.

Z porovnání dat v LINDSEI_CZB, které je patrné v Tabulce 22, se potvrdila Hypotéza 2, tedy že nejčastější VSS kategorií v LINDSEI_CZB jsou předložková slovesa. Ačkoliv rozdíl

mezi typy předložkových a frázových sloves (tj. 80 a 77) není statisticky významný při $p<0.05$, rozdíl mezi tokeny těchto dvou kategorií (tj. 398 a 191) je statisticky významný při $p<0.0001$. Nerodilí mluvčí tedy ve svých promluvách preferují předložková slovesa.

Tabulka 23 prezentuje nejčastější VVS v obou korpusech. S její pomocí je možné ilustrovat poslední hypotézu, která je založená na předpokladu, že nerodilí mluvčí preferují určitá VSS více a případně v neobvyklých kontextech, než je běžné u rodilých mluvčích.[37]

VSS v LINDSEI_CZB a LOCNECB					
VSS	**mluvčí**	**Tokeny**	**VSS**	**Mluvčí**	**Tokeny**
Předložková slovesa: LINDSEI_CZB			**předložková VSS: LOCNECB**		
Talk about	20 (20)	40	Look at	29 (37)	55
Think about	18 (18)	38	Think about	19 (24)	29
Look at	22 (22)	36	Go into	13 (17)	27
Look like	14 (14)	21	Look like	15 (19)	22
Go for	9 (9)	20	Talk about	14 (18)	20
SOUČET	*83*	*155*	*SOUČET*	*90*	*153*
Frázová VSS: LINDSEI_CZB			**frázová VSS: LOCNECB**		
Find out	11 (11)	18	Go back	18 (23)	50
Come back	12 (12)	17	Come back	22 (28)	42
Go on	7 (7)	11	Go out	14 (18)	27
Go back	6 (6)	9	Go on	17 (22)	26
Turn out	7 (7)	9	**Work out**	12 (15)	18
SOUČET	*43*	*64*	*SOUČET*	*83*	*163*
Předložková frázová VSS: LINDSEI_CZB			**Předložková frázová VSS: LOCNECB**		
Look forward to	7 (7)	12	Get on with	5 (6)	8
Go back to	2 (2)	5	Go back to	6 (8)	7
Come up with	2 (2)	3	Look forward to	5 (6)	5
Get away from	2 (2)	2	Come out with	2 (3)	3
Get back to	1 (1)	2	End up with	2 (3)	3
SOUČET	*14*	*24*	*SOUČET*	*20*	*26*

Tabulka 23 – top pět VSS v každé kategorii v daném korpusu

Předložkové *talk about*, frázové *find out* a předložkové frázové *look forward to* jsou nejvíce slibné příklady VSS V Tabulce 23 (společně s *listen to*: 14 výskytů v LINDSEI_CZB vs. žádný v LOCNECB a *turn into/to*: 11 výskytů vs. 1), které nerodilí mluvčí preferují více než rodilí mluvčí, tj. tato VSS jsou tzv. „frázoví medvídci“ (tj. „phrasal teddy-bears“: Ellis: 2012 a Hasselgård, 2019), ke kterým se nerodilí v této studii upínají více než rodilí mluvčí. Kromě *find out* se jedná zejména o nepoměrnou frekvenci výskytu těchto sloves, tj. např. rozdíl ve frekvenci *talk about* v korpusech je 20 tokenů, rozdíl u *look forward to* je sedm tokenů. V případě *find out* se nejedná pouze o zvýšenou frekvenci, ale i o použití v situacích, která nejsou pro rodilé

[37] Čísla v závorce představují normalizovanou frekvenci mluvčích (na 50 mluvčích), jelikož se analýza skládala z promluv 50 nerodilých mluvčích a 39 rodilých mluvčích angličtiny.

mluvčí obvyklá (např. použití *find out* místo jednoslovného *find*). Třetí hypotéza je tedy potvrzena minimálně jedním, při zvážení pouze frekvencí, pěti VSS: nerodilí mluvčí mají své „frázové medvídky“, ke kterým se upínají více než je běžné u rodilých mluvčích. Výsledky ukazují, že jednotliví studenti (jak rodilí, tak nerodilí) mají svá oblíbená VSS, např. *be into* (čtyři tokeny u jednoho mluvčího) v LINDSEI_CZB nebo *go into* (sedm tokenů u jednoho mluvčího). Zároveň se také ukázalo, že nerodilí mluvčí užívají některá VSS výrazně méně, než je běžné u rodilých mluvčích, např. *work out*, které je na páté pozici u rodilých mluvčích (18 tokenů na 12 mluvčích), nerodilí mluvčí použili jen 4krát (4 tokeny a 4 mluvčí), *go into* (27 výskytů a 13 mluvčích v LOCNECB oproti jednomu výskytu v LINDSEI_CZB) a *get on with* (8 výskytů a 5 mluvčích v LOCNECB a dva v LINDSEI_CZB).

Dále se také ukázalo, v případě všech VSS, ale nejvýrazněji u předložkových sloves, že je nutné zvážit roli idiomatičnosti v klasifikaci VSS. Z analýzy vyplývá, že idiomatičnost není binární kategorií, kdy VSS buď je nebo není idiomatické, avšak zdá se být spíše škálou, na které se VSS pohybují, tj. od spíše ne-idiomatických VSS k výrazně idiomatickým VSS. Příkladem jsou VSS *listen to, spy on*, or *thank for* na škále blížící se transparentnosti na jedné straně a *go for (~choose), look after (~care),* or *hit on (~seduce*) na straně druhé, u kterých je idiomatičnost jedním z hlavních rysů. Zdá se, že nejdůležitějším faktorem při klasifikaci předložkových sloves je kolokabilita jejich konstituentů. Frázová slovesa se ve většině případů kloní k idiomatické straně škály, avšak je výhodné vzít tuto škálu v potaz při srovnávání frázových sloves jako *come back* na jedné straně a např. *give up (~quit)* nebo *take up* (*~begin)* na straně druhé. Z dat a Tabulky 23 také vyplývá, že *look at* je nejčastější předložkové sloveso (i VSS) v LOCNECB a třetí nejčastější v LINDSEI_CZB (za *think about* a *talk about*). Zároveň je to také nejvíce rozptýlené VSS v textech obou korpusů, tj. 29 (při normalizované frekvenci 37) mluvčích v LOCNECB a 22 mluvčích v LINDSEI_CZB použilo toto VSS alespoň jednou ve své promluvě. Tyto výsledky z velké míry odpovídají zjištění Biber et al. (1999: 416), kteří také označují *look at* jako nejčastější předložkové sloveso.

Poslední částí práce je závěr, který shrnuje výsledky analýz, připomíná limity metodologie i samostatné analýzy a vyjadřuje se k možnostem dalšího výzkumu. Dále shrnuje hypotézy a odpovědi na ně.

8. APPENDIX

The appendix contains full Tables of MWV categories of both the corpora.

A) LINDSEI_CZB

1. PrV in LINDSEI_CZB; PrV in bold are PrV TypeII

PrV in LINDSEI_CZB				
No	**MWV**	**Raw frequency**	**No of speakers**	**Normalized frequency of speakers**
1	Talk about	40	20	20
2	Think about	38	18	18
3	Look at	36	22	22
4	Look like	21	14	14
5	Go for	20	9	9
6	Listen to	14	10	10
7	Talk to	12	11	11
8	Focus on	11	7	7
9	**Say to**	10	8	8
10	Wait for	9	7	7
11	Work on	9	6	6
12	Come from	8	7	7
13	Get into	8	8	8
14	**Turn into**	8	4	4
15	Be into	7	4	4
16	Think of	7	6	6
17	Deal with	6	4	4
18	Apply for	5	4	4
19	Go through	5	5	5
20	Point at	5	5	5
21	**Say about**	5	5	5
22	Speak to	5	4	4
23	**Connect with**	4	4	4
24	Depend on	4	4	4
25	**Do about**	4	4	4
26	Feel like	4	4	4
27	**Give to**	4	3	3
28	Go with	4	2	2
29	**Know about**	4	4	4
30	**Remind of**	4	3	3
31	**Base on**	3	3	3
32	Care for	3	3	3
33	Concentrate on	3	2	2
34	Consist of	3	2	2
35	Divide into	3	3	3
36	Get over	3	2	2
37	Happen to	3	2	2

PrV in LINDSEI_CZB				
No	**MWV**	**Raw frequency**	**No of speakers**	**Normalized frequency of speakers**
38	Laugh at	3	3	3
39	**Turn to**	3	2	2
40	Come across	2	2	2
41	Engage with	2	1	1
42	Get through	2	2	2
43	Hear of	2	2	2
44	Look after	2	2	2
45	Play with	2	2	2
46	Search for	2	1	1
47	Spy on	2	1	1
48	**Switch into**	2	1	1
49	**Accuse of**	1	1	1
50	**Ask for**	1	1	1
51	**Begin with**	1	1	1
52	Belong to	1	1	1
53	Bump into	1	1	1
54	Delve into	1	1	1
55	**Derive from**	1	1	1
56	Dig into	1	1	1
57	Draw on	1	1	1
58	**Explain to**	1	1	1
59	Feast on	1	1	1
60	**Fill with**	1	1	1
61	Frown on	1	1	1
62	Go after	1	1	1
63	Go into	1	1	1
64	Hit on	1	1	1
65	**Change into**	1	1	1
66	Insist on	1	1	1
67	Look for	1	1	1
68	Look into	1	1	1
69	**Make into**	1	1	1
70	**Put into**	1	1	1
71	Reflect on	1	1	1
72	**Show around**	1	1	1
73	**Stick to**	1	1	1
74	Suffer from	1	1	1
75	**Thank for**	1	1	1
76	**Use as**	1	1	1
77	**Use for**	1	1	1
78	Work for	1	1	1
79	Worry about	1	1	1
80	**Write to**	1	1	1
	TOTAL	398	284	284

2. **PhV in LINDSEI_CZB;** PhV in bold are PhV Type II

PhV in LINDSEI_CZB				
No	**MWV**	**Raw frequency**	**No of speakers**	**Normalized frequency of speakers**
1	**Find out**	18	11	11
2	Come back	17	12	12
3	Go on	11	7	7
4	Go back	9	6	6
5	Turn out	9	7	7
6	Go out	7	7	7
7	**Pick up**	7	4	4
8	**Give up**	4	4	4
9	Go over	4	3	3
10	Sit down	4	3	3
11	**Work** out	4	4	4
12	Carry on	3	3	3
13	Get back	3	3	3
14	Grow up	3	3	3
15	Hang out	3	2	2
16	Help out	3	2	2
17	**Show off**	3	3	3
18	Stand up	3	3	3
19	**Bring up**	2	2	2
20	Come up	2	2	2
21	Dress up	2	1	1
22	Drop out	2	2	2
23	**Figure out**	2	1	1
24	**Fill in**	2	2	2
25	**Get** out	2	2	2
26	Go down	2	2	2
27	Go up	2	2	2
28	**Look around**	2	2	2
29	Meet up	2	2	2
30	Move out	2	2	2
31	**Put on**	2	2	2
32	**Set up**	2	2	2
33	Start off	2	1	1
34	**Take up**	2	2	2
35	**Throw out**	2	1	1
36	**Build up**	1	1	1
37	**Burn down**	1	1	1
38	**Call up**	1	1	1

PhV in LINDSEI_CZB				
No	**MWV**	**Raw frequency**	**No of speakers**	**Normalized frequency of speakers**
39	Catch up	1	1	1
40	**Check over**	1	1	1
41	Come on	1	1	1
42	Come over	1	1	1
43	Die out	1	1	1
44	End up	1	1	1
45	Get up	1	1	1
46	Go away	1	1	1
47	**Hand in**	1	1	1
48	**Keep up**	1	1	1
49	Lie down	1	1	1
50	**Light up**	1	1	1
51	Look back	1	1	1
52	**Look up**	1	1	1
53	Make out	1	1	1
54	**Make up**	1	1	1
55	Move away	1	1	1
56	**Open up**	1	1	1
57	**Print out**	1	1	1
58	**Put together**	1	1	1
59	**Put up**	1	1	1
60	**Read out**	1	1	1
61	**Sell out**	1	1	1
62	**Send away**	1	1	1
63	**Set off**	1	1	1
64	Settle down	1	1	1
65	Settle in	1	1	1
66	**Sign up**	1	1	1
67	**Slow down**	1	1	1
68	**Speed up**	1	1	1
69	**Take in**	1	1	1
70	Take on	1	1	1
71	**Think through**	1	1	1
72	**Throw away**	1	1	1
73	**Tidy up**	1	1	1
74	**Tow away**	1	1	1
75	**Turn on**	1	1	1
76	Turn up	1	1	1
77	**Write down**	1	1	1
	TOTAL	191	159	159

3. PPV in LINDSEI_CZB

PPV in LINDSEI_CZB				
No	**MWV**	**Raw frequency**	**No of speakers**	**Normalized frequency of speakers**
1	Look forward to	12	7	7
2	Go back to	5	2	2
3	Come up with	3	2	2
4	Get away from	2	2	2
5	Get back to	2	1	1
6	Get on with	2	2	2
7	Catch up with	1	1	1
8	Come back to	1	1	1
9	End up with	1	1	1
10	Get around to	1	1	1
11	Hold on to	1	1	1
12	Make up for	1	1	1
13	Try out for	1	1	1
	TOTAL	33	23	23

B) LOCNECB

1. PrV in LOCNECB; PrV in bold are PrV Type II

PrV in LOCNECB				
No	**MWV**	**Raw frequency**	**No of speakers**	**Normalized frequency of speakers**
1	Look at	55	29	37
2	Think about	29	19	24
3	Go into	27	13	17
4	Look like	22	15	19
5	Talk about	20	14	18
6	Think of	17	16	21
7	Come from	16	12	15
8	Apply for	13	8	10
9	**Say to**	12	9	12
10	**Write to**	12	5	6
11	Go for	11	9	12
12	Look for	10	6	8
13	Speak to	10	7	9
14	Worry about	10	5	6
15	**Ask for**	9	3	4
16	Depend on	9	8	10
17	Get on	8	6	8

PrV in LOCNECB				
No	**MWV**	**Raw frequency**	**No of speakers**	**Normalized frequency of speakers**
18	**Do about**	7	6	8
19	**Give to**	7	3	4
20	Talk to	7	5	6
21	Go through	6	3	4
22	Happen to	6	6	8
23	Look into	6	4	5
24	Get into	5	5	6
25	**Know about**	5	4	5
26	Take over	5	3	4
27	Feel like	4	4	5
28	Hear of	4	4	5
29	Look through	4	4	5
30	Appeal to	3	3	4
31	Care for	3	1	1
32	Concentrate on	3	3	4
33	Deal with	3	2	3
34	Look after	3	3	4
35	Relate to	3	3	4
36	**Say about**	3	3	4
37	Aim at	2	1	1
38	Be into	2	2	3
39	Come across	2	2	3
40	Cope with	2	2	3
41	Focus on	2	2	3
42	Fuss over	2	1	1
43	Go with	2	2	3
44	Hear from	2	1	1
45	Live for	2	1	1
46	Live on	2	1	1
47	**Make sth into sth**	2	2	3
48	Rely on	2	2	3
49	Abide by	1	1	1
50	**Base on**	1	1	1
51	**Begin with**	1	1	1
52	**Blame for**	1	1	1
53	Bump into	1	1	1
54	Call on	1	1	1
55	Cater for	1	1	1
56	Cling to	1	1	1
57	Comment on	1	1	1

PrV in LOCNECB				
No	**MWV**	**Raw frequency**	**No of speakers**	**Normalized frequency of speakers**
58	Count for	1	1	1
59	Decide on	1	1	1
60	**Explain to**	1	1	1
61	Frown on	1	1	1
62	Get past	1	1	1
63	Get through	1	1	1
64	**Hear about**	1	1	1
65	Look on	1	1	1
66	Point at	1	1	1
67	Remind of	1	1	1
68	Run on	1	1	1
69	Save on	1	1	1
70	**Turn into**	1	1	1
71	Wait for	1	1	1
72	Work for	1	1	1
73	Work on	1	1	1
	TOTAL	426	268	381

2. **PhV in LOCNECB;** PhV in bold are of Type II

PhV in LOCNECB				
No	**MWV**	**Raw frequency**	**No of speakers**	**Normalized frequency of speakers**
1	Go back	50	18	23
2	Come back	42	22	28
3	Go out	27	14	18
4	Go on	26	17	22
5	**Work out**	18	12	15
6	Get out	17	14	18
7	Get up	16	10	13
8	End up	14	11	14
9	Get back	14	10	13
10	Come up	13	12	15
11	Come in	12	9	12
12	Go away	12	11	14
13	Sit down	12	10	13
14	Go over	11	6	8
15	Get away	9	8	10
16	**Pick up**	9	6	8
17	Stand up	9	3	4
18	**Take out**	9	7	9

PhV in LOCNECB				
No	**MWV**	**Raw frequency**	**No of speakers**	**Normalized frequency of speakers**
19	Wake up	9	5	6
20	Carry on	8	7	9
21	**Find out**	8	8	10
22	Get in	8	7	9
23	Go off	8	6	8
24	Move away	8	4	5
25	**Put on**	8	7	9
26	**Show off**	8	6	8
27	Turn out	8	6	8
28	Get together	7	5	6
29	**Set up**	7	6	8
30	**Start off**	7	5	6
31	**Take off**	7	6	8
32	Come on	6	4	5
33	Move out	6	4	5
34	**Put together**	6	5	6
35	**Sort out**	6	5	6
36	Come over	5	5	6
37	Come round	5	4	5
38	Fit in	5	5	6
39	**Beat up**	4	3	4
40	**Bring i**	4	3	4
41	**Build up**	4	4	5
42	Come through	4	2	3
43	**Cut off**	4	3	4
44	Grow up	4	3	4
45	Look around	4	4	5
46	**Make up**	4	3	4
47	**Put off**	4	4	5
48	Stop off	4	2	3
49	**Take on**	4	2	3
50	Come along	3	3	4
51	Come out	3	2	3
52	**Curl up**	3	2	3
53	**Fill in**	3	3	4
54	**Give up**	3	3	4
55	Go ahead	3	3	4
56	Go along	3	3	4
57	**Hand in**	3	3	4
58	Help out	3	1	1

PhV in LOCNECB				
No	**MWV**	**Raw frequency**	**No of speakers**	**Normalized frequency of speakers**
59	Look back	3	3	4
60	Miss out	3	2	3
61	**Pay back**	3	1	1
62	**Pick out**	3	2	3
63	**Point out**	3	2	3
64	**Rip off**	3	2	3
65	**Set off**	3	3	4
66	Shut up	3	2	3
67	Add up	2	2	3
68	**Back up**	2	2	3
69	**Bring up**	2	2	3
70	Burst out	2	2	3
71	**Close off**	2	2	3
72	Come together	2	2	3
73	**Cover up**	2	1	1
74	**Eat out**	2	1	1
75	**Fight back**	2	1	1
76	Get by	2	2	3
77	Give back	2	1	1
78	**Give out**	2	2	3
79	Go forward	2	1	1
80	Go up	2	2	3
81	**Hand out**	2	2	3
82	Head off	2	1	1
83	**Heat up**	2	1	1
84	Keep on	2	2	3
85	**Keep up**	2	2	3
86	Lie down	2	2	3
87	**Lock up**	2	1	1
88	Meet up	2	2	3
89	**Mix up**	2	1	1
90	Move on	2	2	3
91	**Put sth up**	2	2	3
92	**Ring up**	2	2	3
93	Run out	2	2	3
94	**Send off**	2	2	3
95	**Shut down**	2	2	3
96	**Split up**	2	1	1
97	Stand out	2	1	1
98	**Start up**	2	2	3

PhV in LOCNECB				
No	**MWV**	**Raw frequency**	**No of speakers**	**Normalized frequency of speakers**
99	Stay up	2	2	3
100	**Take up**	2	2	3
101	**Turn off**	2	2	3
102	Work up	2	1	1
103	Write off	2	2	3
104	**Act out**	1	1	1
105	**Ask sb out**	1	1	1
106	**Belt sth out**	1	1	1
107	**Block sb up**	1	1	1
108	**Bolster up**	1	1	1
109	Break up	1	1	1
110	**Bring down**	1	1	1
111	**Chunk out**	1	1	1
112	Come off	1	1	1
113	Crop up	1	1	1
114	Cut down	1	1	1
115	**Cut up**	1	1	1
116	Do up	1	1	1
117	Dress up	1	1	1
118	Fade out	1	1	1
119	Fall out	1	1	1
120	Figure out	1	1	1
121	**Fill out**	1	1	1
122	**Fill up**	1	1	1
123	**Hold back**	1	1	1
124	**Hurry up**	1	1	1
125	Kick off	1	1	1
126	**Leave out**	1	1	1
127	**Light up**	1	1	1
128	**Link up**	1	1	1
129	**Look up**	1	1	1
130	**Make sth out**	1	1	1
131	Move up	1	1	1
132	Open up	1	1	1
133	**Pack up**	1	1	1
134	**Pay off**	1	1	1
135	**Phone up**	1	1	1
136	**Plan out**	1	1	1
137	Pop off	1	1	1
138	Pour down	1	1	1

PhV in LOCNECB				
No	**MWV**	**Raw frequency**	**No of speakers**	**Normalized frequency of speakers**
139	**Push down**	1	1	1
140	Queue up	1	1	1
141	Roll over	1	1	1
142	**Save up**	1	1	1
143	**Sell out**	1	1	1
144	Settle in	1	1	1
145	**Sign up**	1	1	1
146	Sit up	1	1	1
147	Slip up	1	1	1
148	**Spread out**	1	1	1
149	Stay out	1	1	1
150	**Sum up**	1	1	1
151	**Take back**	1	1	1
152	**Throw out**	1	1	1
153	Turf out	1	1	1
154	**Turn down**	1	1	1
155	Write back	1	1	1
	TOTAL	681	497	637

3. PPV in LOCNECB

Phrasal-prepositional verbs – LOCNECB				
No	**MWV**	**Raw frequency**	**No of speakers**	**Normalized frequency of speakers**
1	Get on with	8	5	6
2	Go back to	7	6	8
3	Look forward to	5	5	6
4	Come out with	3	2	3
5	End up with	3	2	3
6	Fit in with	3	2	3
7	Stand up to	3	1	1
8	Come out of	2	1	1
9	Come up with	2	2	3
10	Make up for	2	2	3
11	Get away with	2	2	3
12	Keep up with	1	1	1
13	Look out for	1	1	1
14	Put up with	1	1	1
15	Stay out of	1	1	1
16	Get through to	1	1	1
17	Get down to	1	1	1
	TOTAL	46	36	46